NATURAL VARIATION IN SENSITIVITY TO A LOSS OF CHLOROPLAST TRANSLATION IN ARABIDOPSIS

By

NICOLE BRYANT PARKER

Bachelor of Science in Botany
Oklahoma State University
Stillwater, OK
2011

Bachelor of Science in Microbiology and
Molecular Genetics
Oklahoma State University
Stillwater, OK
2011

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 2017

NATURAL VARIATION IN SENSITIVITY TO A LOSS OF CHLOROPLAST TRANSLATION IN ARABIDOPSIS

Dissertation Approved:

Dr. David Meinke

Dissertation Adviser

Dr. Andrew Doust

Dr. Mark Fishbein

Dr. Gerald Schoenknecht

Dr. Robert Burnap

Name: NICOLE BRYANT PARKER

Date of Degree: DECEMBER, 2017

Title of Study: NATURAL VARIATION IN SENSITIVITY TO A LOSS OF CHLOROPLAST TRANSLATION IN ARABIDOPSIS

Major Field: PLANT SCIENCE

Abstract:

This dissertation describes my role in an NSF-funded research project in the Meinke laboratory that began as a natural variation study and genetic analysis to uncover the nuclear genes involved in the differing responses of plant species to a loss of chloroplast translation. To identify these nuclear genes, we analyzed 152 natural accessions of Arabidopsis (*Arabidopsis thaliana*) on spectinomycin, an inhibitor of chloroplast translation, and crossed wild-type plants of the tolerant Tsu-0 accession with plants segregating for an embryo-defective (*emb*) mutation that eliminated chloroplast translation in the sensitive "Nossen" accession. Through this study, we found a single suppressor locus (*ACC2*), an enhancer of the suppressor, and additional modifiers that further increase embryo development. After determining that *ACC2* suppresses the loss of chloroplast translation in *emb* mutants, we expanded our project to include a detailed analysis of defects in *ACC2* and the consequences of various mutations on a class of proteins essential for growth and development in plants. Remarkably, some of the most sensitive accessions contain null alleles of *ACC2*, including "Nossen". For the final part of my role in this project, I focused on using a candidate gene approach to identify additional genetic modifiers of this system. Overall, the project described throughout this dissertation utilized natural variation in Arabidopsis accessions to study the effects of mutations, especially deleterious mutations, on a protein (ACCase) that is essential for fatty acid biosynthesis in eukaryotes. We also developed an understanding of some of the mechanisms behind the diverse phenotypic responses plant species have when translation of the chloroplast genome is blocked. Furthermore, our identification of accessions hypersensitive to spectinomycin has led to a more efficient method for plastid transformation in Arabidopsis (Yu et al., 2017).

TABLE OF CONTENTS

Chapter Page

Chapter Page

Chapter Page

LIST OF TABLES

Table Page

Table Page

LIST OF FIGURES

Figure Page

Figure Page

CHAPTER I

INTRODUCTION

Arabidopsis thaliana is a Model System for Plant Biology

The use of Arabidopsis (*Arabidopsis thaliana*) as a model system for plant biology began in the 1940s with Friedrich Laibach, who noted that in contrast to many agricultural plants, Arabidopsis grows rapidly in small spaces, produces a large number of offspring, and contains a small number of chromosomes (Laibach, 1943; Sommerville and Koornneef, 2002). Shortly after Laibach's publication, a small but active Arabidopsis research community was built, which continued through the 1960s. Early research included the analysis of induced mutations generated through X-irradiation, natural variation studies of seed dormancy and flowering time, and chemical mutagenesis studies looking at embryo-lethal mutants (Rédei, 1970). One important advance was the formation of the Arabidopsis Information Service (AIS) newsletter and seed stock center in Germany (Meyerowitz, 2001). The early 1970s brought a decline of research in Arabidopsis and an increased interest in other systems, including petunia and tobacco, where plants could be regenerated more readily from cells in culture (Meinke et al., 1998; Koornneef and Meinke, 2010).

Interest in Arabidopsis research was revitalized in the late 1970s and early 1980s, when plant biologists were seeking a model organism for molecular genetics. George Rédei at the University of Missouri published an important review on Arabidopsis (Rédei, 1975) that soon increased research in the field (Sommerville and Koornneef, 2002; Koornneef and Meinke, 2010). Subsequent publications included the work of Meinke and Sussex (1979a,b) on the use of Arabidopsis embryo-lethal mutants to study plant embryo development, Sommerville and Ogren (1980) on mutants altered in photorespiration, and Koorneef et al. (1983) establishing the first comprehensive genetic map. Several publications in the mid-1980s described the advantage of Arabidopsis' small genome in the field of molecular genetics (Leutwiler et al., 1984; Meyerowitz and Pruitt, 1985). The advantages of Arabidopsis as a model genetic organism attracted the interest of relatively young plant biologists, and brought scientists working on other model systems to the field of Arabidopsis research.

The 1980s and 1990s saw the Arabidopsis community continue to flourish, and the establishment of many useful research tools. One of the most important tools was effective transformation procedures, which allowed Arabidopsis researchers to analyze gene expression patterns and develop large collections of transfer-DNA (T-DNA) insertion mutants (Feldmann and Marks, 1987; Bechtold et al., 1993; Clough and Bent, 1998; Alonso et al., 2003). The development of a genetic model for floral morphogenesis (Weigel and Meyerowitz, 1994) expanded the Arabidopsis field by illustrating how genetic approaches could be applied to complex biological processes (Sommerville and Koornneef, 2002; Koornneef and Meinke, 2010). The genetic map was soon expanded using molecular markers such as restriction fragment length polymorphisms (RFLPs; Chang et al., 1988), simple sequence length polymorphisms (SSLPs; Bell and Ecker, 1994), cleaved amplified polymorphic sequences (CAPSs; Konieczny and Ausubel, 1993), and amplified fragment length polymorphisms (AFLPs; Alonso-Blanco et al., 1998). An updated classical map, published in 1998, contained 462 genes dispersed across all five chromosomes (Meinke et al., 1998; Koornneef and Meinke, 2010). Once the sequence of the Arabidopsis genome was published, efforts were made

to integrate the classical genetic map with the sequence-based physical map using genes with known mutant phenotypes (Meinke et al., 2003; Meinke et al., 2009). The most current map of genes with mutant phenotypes contains 2,400 loci, about 9% of the total number of genes in the Arabidopsis genome (Lloyd and Meinke, 2012).

Over the past 20 years, three large collaborative projects have significantly advanced the field of Arabidopsis research. The first project, the completion of the Arabidopsis genome sequence (AGI, 2000), paved the way for Arabidopsis research to expand into the age of genomics (Sommerville and Koornneef, 2002; Koornneef and Meinke, 2010). The Arabidopsis 2010 project soon followed, with a focus on understanding the functions of all 25,000 protein-coding genes (Chory et al., 2000; Sommerville and Koornneef, 2002; Koornneef and Meinke, 2010). The 1001 Genomes Project, begun in 2008 and still ongoing, involves whole-genome sequencing of Arabidopsis accessions, with the goal of enhancing research linking phenotypes to genotypes (http://1001genomes.org/). To date, 1,135 different accessions have been sequenced (The 1001 Genomes Consortium, 2016).

To further collaboration, the Arabidopsis community has developed numerous shared resources. These include two centralized databases of genetic and molecular data: The Arabidopsis Information Resource (TAIR) at www.arabidopsis.org (Rhee et al., 2003), and the Arabidopsis Information Portal (ARAPORT) at www.araport.org (Cheng et al., 2017). Seed stock centers are another example of shared resources. Thirty years ago, the AIS newsletter published a list of around 1,000 available stocks, mainly natural accessions and a limited number of mutant lines (Provart et al, 2015). Two major stock centers were founded in the early 1990s: the Nottingham Arabidopsis Stock Centre (NASC) and the Arabidopsis Biological Resource Center (ABRC). These centers now contain more than 900,000 stocks (Koornneef and Meinke, 2010; Provart et al, 2015).

Embryo-Defective Mutants of Arabidopsis Have Been Characterized in Detail

Over the past 35 years, the Meinke laboratory at Oklahoma State University has isolated and characterized several thousand embryo-defective (*emb*) mutants of Arabidopsis and catalogued large numbers of essential genes required for seed and embryo development. The methods used to identify these mutants were first described by Müller (1963). Later, Meinke and Sussex (1979a,b) discussed the benefits of using *emb* mutants in research on plant embryo development. Before the era of sequencing, *emb* mutants were identified through forward genetic screens of mutant plants produced using chemical mutagens such as ethyl-methanesulfonate (EMS). After treatment, plants were screened for embryo lethality (25% aborted seeds), and mutants segregating as Mendelian recessives were isolated and characterized (Meinke and Sussex, 1979a,b). The development of T-DNA insertional mutagenesis enabled large-scale screens of mutants that allowed for quicker identification of the disrupted gene through amplification of sequences flanking the T-DNA insertion site (McElver et al., 2001; Meinke, 2008; Meinke, 2013). T- DNA insertion mutants were also used in a reverse genetic approach to characterize *emb* mutants disrupted in known genes believed to be essential (Meinke, 2013).

Of the estimated 750-1,000 *EMB* genes in the Arabidopsis genome, over 400 have been cloned and sequenced to date (Muralla et al., 2011). The SeedGenes project (http://seedgenes.org) was established in 2002 to create a centralized database containing information on loss-of-function mutant alleles that give rise to a seed or embryo phenotype (Tzafrir et al., 2003). The current database, updated in December 2010, contains 888 mutant alleles and 481 genes (Meinke et al., 2013). The *emb* mutant alleles in SeedGenes have been placed into six categories based on their terminal embryo phenotype: preglobular, preglobular/globular, globular, transition, cotyledon, or unresolved. The essential genes listed in the database have been divided into three classes (Muralla et al., 2011): (1) embryo defective, characterized by defects in seed development; (2) seed pigment, characterized by defects in seed pigmentation; and (3) 50% defective seeds, characterized by

approximately 50% mutant seeds in selfed heterozygotes. The phenotypes of mutant alleles have been examined in considerable detail (Meinke et al., 2008; Muralla et al., 2011).

Of the 400 identified *EMB* genes, 119 are predicted to encode chloroplast-localized proteins. These genes can be divided further into three groups based on protein function: (1) proteins involved in the biosynthesis of metabolites such as amino acids and vitamins; (2) proteins associated with import, modification, and localization of proteins within the chloroplast; and (3) proteins required for translation of RNAs encoded by the chloroplast genome (Bryant et al., 2011). This third category is most relevant to the project described in this dissertation. Around 23% of chloroplast-localized EMB proteins are involved in chloroplast translation, including plastid ribosomal proteins (PRPs), chloroplast-localized aminoacyl-tRNA synthetases (AARSs), and chloroplast-localized pentatricopeptide repeat (PPR) proteins, which function in RNA binding and modification (Berg et al., 2005; Schmitz-Linneweber and Small, 2008; Bryant et al., 2011, Romani et al., 2012; Tiller and Bock, 2014). Chloroplast translation is therefore required for embryo development in Arabidopsis. Mutations in genes that encode proteins of the photosynthetic machinery lead to reduced pigmentation in the embryo rather than lethality (Bryant et al., 2011). These observations raise an important question related to this project: What specific protein(s) encoded by the chloroplast genome are required (must be translated from chloroplast-encoded mRNAs) for embryo development in Arabidopsis?

The Chloroplast Genome in Arabidopsis Contains Essential Genes

The complete nucleotide sequence of the chloroplast genome in Arabidopsis was published in 1999 (Sato et al., 1999). The chloroplast genome contains 128 genes: four encoding ribosomal RNAs (rRNAs), 37 encoding transfer RNAs (tRNAs), and 87 encoding proteins (Figure 1). The protein-coding genes can be divided into five categories based on protein function: transcription, translation,

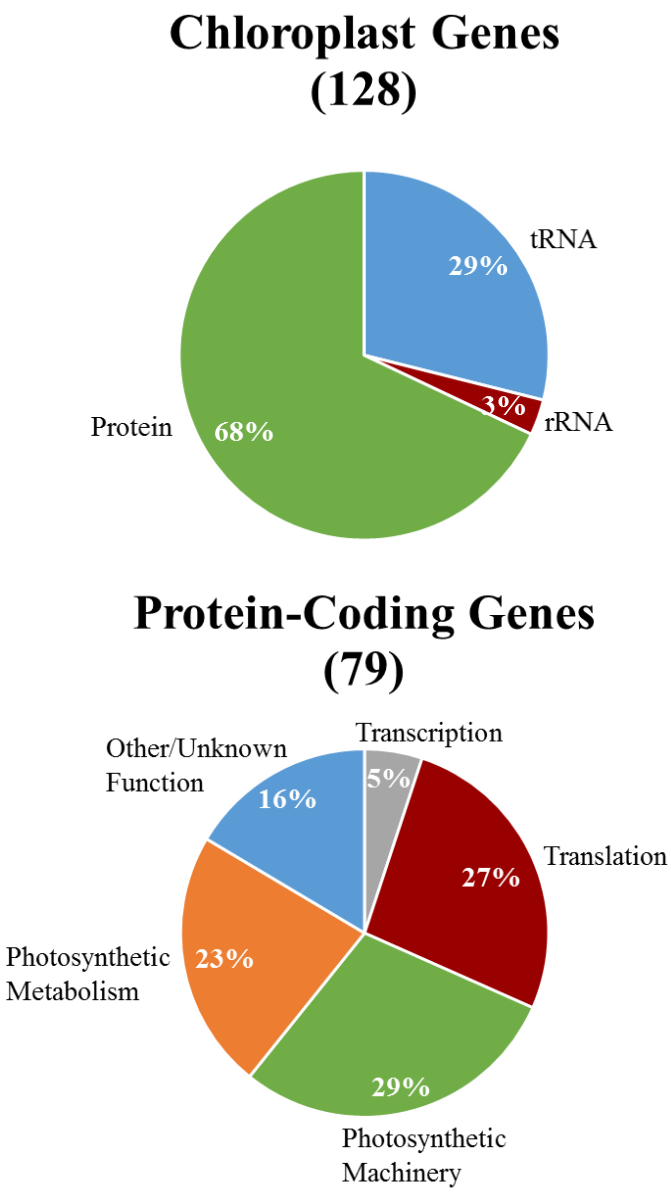


Figure 1. Distribution of Genes in the Chloroplast Genome of Arabidopsis. (A) The chloroplast genome of Arabidopsis contains 128 genes, which can be divided into three classes: protein-coding genes, rRNA genes, and tRNA genes. (B) The protein-coding genes on can be further divided into five categories based on protein function. The information shown in this figure was taken from Sato et al. (1999).

photosynthetic machinery, photosynthetic metabolism, and other/unknown function (Figure 1; Sato et al., 1999). In addition to genes encoding proteins involved in gene expression and photosynthesis, 11 genes encode proteins with other functions. Four of these, *hypothetical chloroplast open reading frame 3* (*ycf3*), *ycf4*, *ycf5* and *ycf6*, are involved in protein assembly and stability of the photosynthetic machinery (Hager et al., 1999; Naver et al., 2001; Goddard et al., 2010; Krech et al., 2012). A fifth gene, *ycf9*, encodes PsbZ, a subunit of the Photosystem II complex (Swiatek et al., 2001; Tang et al., 2016). *Ycf10* is believed to function in efficient transport of inorganic carbon across the chloroplast membrane (Rolland et al., 1997). *MatK* encodes a maturase protein involved in RNA splicing of type II introns within the chloroplast (Vogel et al., 1997; Uchoi et al., 2016).

Four chloroplast genes have been identified as essential by targeted gene disruptions in tobacco (*Nicotiana tabacum*). *Acetyl-coenzyme A carboxylase D* (*accD*) was shown to be required for leaf development, *caseinolytic protease P1* (*clpP1*) essential for shoot development, and *ycf1* and *ycf2* required for cell survival (Drescher et al., 2000; Kuroda and Maliga, 2003; Kode et al., 2005). The *clpP1* gene encodes a subunit of a chloroplast-localized protease complex known to be required for chloroplast function in Arabidopsis (Ramos-Vega et al., 2015). Kikuchi et al. (2013) discovered that *ycf1* encodes a component (Translocon at Inner envelope membrane of the Chloroplast; Tic214) of the TIC chloroplast protein import system located at the inner envelope membrane of chloroplasts. *Ycf2* is believed to also function in chloroplast protein import (Parker et al., 2016; Masato Nakai, personal communication).

The *accD* gene encodes the β-carboxyl transferase subunit of the heteromeric acetyl-coenzyme A carboxylase (ACCase), which is localized to the chloroplast. The other subunits of this enzyme are nuclear-encoded proteins, of which one (*Chloroplastic Acetyl-Coenzyme A Carboxylase*; *CAC1A*) is known to be required for embryo development (Li et al., 2011). This protein functions during the early stages of fatty acid biosynthesis within the chloroplast to catalyze the conversion of acetyl-CoA to malonyl-CoA (Ohlrogge and Browse, 1995; Li et al., 2011). The project described here

focused on *accD* as the essential chloroplast gene most important (rate limiting) for seedling and embryo development in Arabidopsis.

Plant Species Differ in Response to a Loss of Chloroplast Translation

Not all plant species are equally sensitive to a loss of chloroplast translation. Zubko and Day (1998) exposed seedlings to spectinomycin, an inhibitor of chloroplast translation, and found that Brassica (*Brassica napus*) was more tolerant than tobacco, which was more sensitive than Arabidopsis. The reason for these differences remained unknown. The response of grass species to the loss of chloroplast translation is more complicated. In the late 1970s and early 1980s, albino leaf regions in mutants of maize (*Zea mays*) and barley (*Hordeum vulgare*) were shown to lack chloroplast ribosomes (Walbot and Coe, 1979; Siemenroth et al., 1981). Around 25 years later, Asakura and Barkan's (2006) work on splicing mutant homologs further showed that maize plants could tolerate a loss of chloroplast translation through the loss of a single chloroplast-localized splicing factor, *caf2*. They also showed that a null allele of an orthologous protein in Arabidopsis, *Atcaf2*, was embryo-lethal (Asakura and Barkan, 2006).

In contrast to these discoveries, multiple maize mutants disrupted in chloroplast translation have been shown to exhibit embryo lethality. Unlike the aborted seeds in Arabidopsis *emb* mutants, the seeds of these maize mutants have normal development of the endosperm tissue (Ma and Dooner, 2004; Magnard et al., 2004; Sosso et al., 2012; Zhang et al., 2013; Shen et al., 2013; Li et al., 2015). Zhang et al. (2013) have shown that the effects of disrupting a gene essential for chloroplast translation in maize is dependent on the genetic background. Using mutations in the maize *Why1* gene, which is believed to be involved in stability of the chloroplast genome and the formation of chloroplast ribosomes, they showed that mutants in the *W22* background were embryo defective while mutants in the *B73* and *Mo17* backgrounds grew as albino seedlings (Zhang et al., 2013). The

current hypothesis for the background effect in maize is the differing activity of retrograde signaling pathways between the chloroplast and nuclear genomes. When translation of the chloroplast genome is disrupted in backgrounds like *W22*, a signal is thought to be sent out to terminate cell activity within the embryo, which eventually leads to embryo lethality (Terry and Smith, 2013; Zhang et al., 2013; Li et al., 2015).

During the evolution of the Poaceae family, *accD* was lost from the chloroplast genome along with *ycf1* and *ycf2*. Loss of *accD* means the absence of the heteromeric ACCase protein, which catalyzes a crucial step in fatty acid biosynthesis. Grasses have compensated for this loss with a nuclear-encoded, homomeric ACCase that is targeted to the chloroplast (Maier et al., 1995; Jansen et al., 2007; Guisinger et al., 2010). Members of the Brassicaceae also have a nuclear-encoded, homomeric ACCase that is targeted to the chloroplast (Schulte et al., 1997; Babiychuk et al., 2011). This novel gene, *ACC2*, arose during the evolution of the Brassicaceae family from a duplication of *ACC1*, a homomeric, cytosolic ACCase that is involved in later stages of fatty acid biosynthesis. In Brassica and Arabidopsis, *ACC2* is targeted to the chloroplast, where it can partially compensate for the loss of the heteromeric ACCase when chloroplast translation is blocked. However, *ACC2* is poorly expressed in the Columbia accession of Arabidopsis. A model of this mechanism of partial nuclear compensation for a loss of heteromeric, chloroplast-encoded ACCase in Brassicaceae is shown in Figure 2. The project described in this dissertation used natural variation in Arabidopsis accessions to study this nuclear compensation pathway and the effects of *acc2* mutations on plant growth and development in the absence of chloroplast translation.

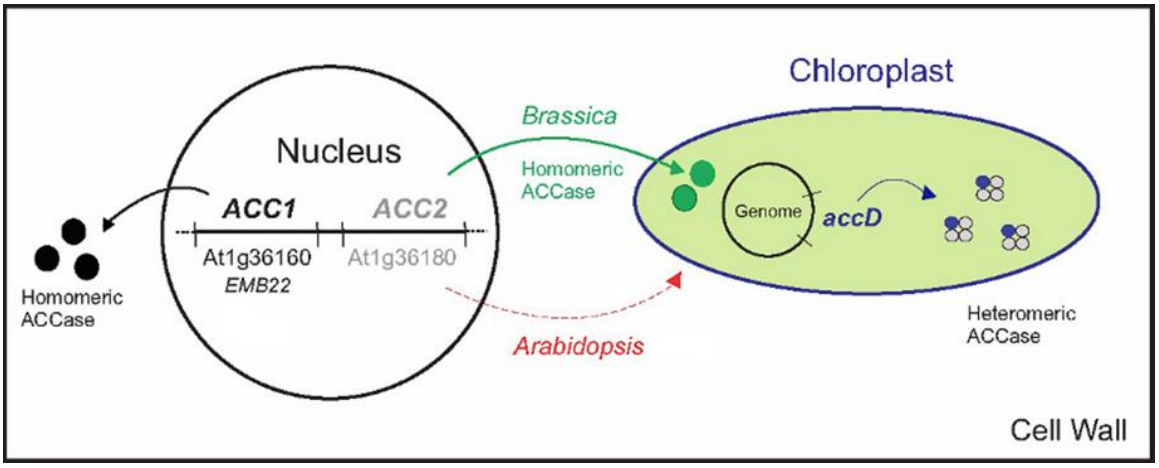


Figure 2. Nuclear Compensation for Loss of Heteromeric ACCase in Brassicaceae. In Brassica, *ACC2*, a duplicated ACCase gene in the nuclear genome, is transcribed and localized to the chloroplast where it can compensate for the loss of *accD* when chloroplast translation is blocked. In the Arabidopsis Columbia accession, *ACC2* is poorly expressed, and there is more limited compensation for the loss of *accD*. Adapted from a drawing by Rosanna Muralla.

Outline and Scope of Dissertation

This dissertation describes my role in an NSF-funded research project in the Meinke laboratory that utilized natural variation in Arabidopsis to study why plant species differ in their ability to tolerate a loss of chloroplast translation. I began working in the Meinke laboratory as an undergraduate researcher, assisting with the characterization of *EMB* genes predicted to encode chloroplast-localized proteins. The work from that project, which was published in 2011, led us to the question: What protein(s) encoded by the chloroplast genome are required for embryo development in Arabidopsis? (Bryant et al., 2011). The project described here began as a natural variation study and genetic analysis to uncover the nuclear genes involved in the differing responses of plant species to a loss of chloroplast translation. This was later expanded to include a detailed analysis of defects in *ACC2* and the consequences of various mutations on a class of proteins essential for growth and development in plants. Two articles have been published in Plant Physiology describing the results of this project (Parker et al., 2014; 2016). The work published in Parker et al. (2014) is described in Chapters 3 and 4 of this dissertation, and that published in Parker et al. (2016) is described in Chapters 3 and 5.

The work on this project was divided between three people: Dr. David Meinke (DM), Dr. Yixing Wang (YW), and myself (NP). DM conceived and managed this project, performed some of the crosses, helped substantially with the embryo phenotyping, and wrote the two articles in Plant Physiology, with input from YW and NP. YW designed and completed all of the molecular biology experiments for this project, with the exception of the candidate gene approach described in Chapter 6. NP screened the phenotypes of all seedlings grown on spectinomycin and lincomycin, completed a significant amount of the embryo phenotyping with a focus on lines with the most advanced embryos, performed many of the crosses, handled all of the ACCase sequence alignments, maintained plants and seed stocks, and carried out the candidate gene approach discussed in Chapter 6.

Looking ahead in this dissertation, the second chapter is a literature review of selected topics related to this project. The third chapter describes a detailed study of the natural variation observed among Arabidopsis accessions in response to a loss of chloroplast translation. This work utilized spectinomycin, an antibiotic known to inhibit translation of the chloroplast genome. The fourth chapter focuses on the use of *emb* mutants to identify factors that enhance tolerance to a loss of chloroplast translation, including a suppressor of early embryo arrest, an enhancer of the suppressor, and additional modifiers to the system. The fifth chapter describes the analysis of missense mutations in *ACC1* and *ACC2* found in natural accessions of Arabidopsis. The sixth and final chapter describes a candidate gene approach that was used in an attempt to identify other factors that increase tolerance to a loss of chloroplast translation.

CHAPTER II

REVIEW OF LITERATURE

Chloroplast Genome Content Across Plant Species

Chloroplast genomes of higher plants are comprised of circular, double-stranded DNA that is highly conserved in size, structure, and gene content. Since the first fully-sequenced chloroplast genome from tobacco was published in 1986 (Shinozaki et al., 1986), the number of sequenced chloroplast genomes has grown exponentially (Curci et al., 2016; Daniell et al., 2016). Currently, over 1,500 chloroplast sequences are available in the National Center for Biotechnology Information (NCBI) Genome Database (https://www.ncbi.nlm.nih.gov/genome/browse/). In higher plants and algae, the chloroplast genome, on average, is around 120-160 kb in length, contains roughly 130 genes, and is composed of four distinct regions: two inverted repeats (IR), a large single-copy region (LSC), and a small single-copy region (SSC) (Olmstead and Palmer, 1994; Jansen et al., 2005; Odintsova and Yurina, 2005; Curci et al., 2016). Most genes are found within the LSC and SSC regions, with the exception of a handful of rRNA and tRNA genes that are located in the IR regions (Odintsova and Yurina, 2005; Chumley et al., 2006).

Comparisons of sequenced chloroplast genomes have been used to study the evolution of higher plants through the rates of nucleotide substitutions, gene insertion/deletions, and genomic rearrangements (Jansen et al., 2007). Rates of synonymous nucleotide substitutions, also known as silent mutations, in chloroplast genomes have been shown to be around half of those found in plant nuclear genomes, with even lower rates in IR regions (Wolfe et al., 1987; Odintsova and Yurina, 2005; Daniell et al., 2016). Large deletions and inversions of the chloroplast genome have been documented in a number of species, and used to resolve phylogenetic relationships (Olmstead and Palmer, 1994; Jansen et al., 2008). For example, all members of the IR lacking clade (IRLC) of legumes have completely lost one copy of the IR region (Palmer and Thompson, 1982; Jansen et al., 2008), and conifers have seen widespread reduction or deletion of their IR regions (Raubeson and Jansen, 1992; Lin et al., 2010). Not surprisingly, most photosynthetic genes have been lost from non-photosynthetic, parasitic plants (dePamphilis and Palmer, 1990).

Along with gene losses, a number of plant species contain a non-functional copy of one or more genes in the chloroplast genome. These pseudogenes can be identified using bioinformatics tools to scan the chloroplast genome for regions that are similar to known chloroplast genes but lack an entire open reading frame (Logacheva et al., 2011). The plastid genes most often retained (*accD, ycf1, ycf2, clpP1*) are those shown experimentally to be essential in tobacco. While still rare, individual gene losses in the chloroplast genome are more common than large deletions and inversions. These gene losses have been studied extensively throughout higher plants, especially in the context of phylogenetic analyses (Jansen et al., 2007). In some cases, chloroplast gene loss is accompanied by the appearance of a compensatory gene in the nuclear genome (Li et al., 2016; Liu et al., 2016). In plant species where *accD* is non-functional or has been lost from the chloroplast genome, there is likely either a duplication of the homomeric ACCase in the nucleus, which occurs in maize (Jansen et al., 2007; Guisinger et al., 2010), or *accD* itself has incorporated into the nuclear genome and is targeted back to the

chloroplast, which occurs in *Trachelium caeruleum* (Harberle et al., 2008; Rousseau-Gueutin et al., 2013). Table 1 lists several losses and pseudogenization of genes related to this project: *accD*, *ycf1*, *ycf2*, *clpP1*.

Spectinomycin and Other Inhibitors of Chloroplast Translation

Antibiotics are used in plant biology research to block translation of the chloroplast genome through a variety of different mechanisms. Several antibiotics and their modes of action are described in this section, including spectinomycin, which was the main antibiotic used in this project, and lincomycin, which was used to confirm results from the spectinomycin studies (Parker et al., 2014, 2016). Other antibiotics with different modes of action noted here include streptomycin, tetracycline, and pactamycin.

Spectinomycin inhibits translation by binding to the 30S ribosomal subunit and interfering with peptidyl-tRNA translocation from the A-site to the P-site in the ribosome (Carter et al., 2000). Specifically, spectinomycin binds within the minor groove of helix 34 of the 16S rRNA in the head of the 30S ribosomal subunit, where it stabilizes the helix during the elongation cycle in translation (Johanson and Hughes, 1995; Carter et al., 2000). Mutagenesis studies on plants resistant to spectinomycin have been used to identify specific nucleotides in the 16S rRNA that interact with the antibiotic. These studies have shown that spectinomycin binding is sequence specific within helix 34 (Carter et al., 2000; Wirmer and Westhof, 2006; Dudas et al., 2012). Resistance to spectinomycin can also be found in plants that have mutations in the S5 ribosomal protein, which is located next to helix 34 and believed to stabilize this region of the 30S ribosomal subunit (Carter et al., 2000; Wirmer and Westhof, 2006).

Lincomycin is a member of the lincosamide class of antibiotics. Members within this

Table 1. Lineages Missing Essential Genes from the Chloroplast Genome.

Organism	Lineages Missing Essential Genes [a]				Reference
	accD	*ycf1*	*ycf2*	*clpP1*	
Acorus	X	Present	Present	Present	Jansen et al., 2007
Aristolochia	Present	PS	Present	Present	Zhou et al., 2017
Asclepias	PS	PS	Present	PS	Straub et al., 2011
Campanula	PS	Present	Present	Present	Rousseau-Gueutin et al., 2013
Cynodon	X	PS	PS	Present [b]	Huang et al., 2017
Epimedium	Present	Present	Present	PS	Sun et al., 2016
Gentiana	Present	X	Present	Present	Fu et al., 2016
Grasses (Six crops)	X	X	X	Present	Jansen et al., 2007; Guisinger et al., 2010
Jasminum	X	Present	Present	Present	Jansen et al., 2007
Pelargonium	X	Present	Present	Present	Jansen et al., 2007
Passiflora	X	X	Present	X	Jansen et al., 2007
Primula	PS	Present	Present	Present	Liu et al., 2016
Scaevola	Present	Present	Present	X	Jansen et al., 2007
Sciadopitys	X	Present	Present	Present	Li et al., 2016
Trachelium	PS	X	X	X	Jansen et al., 2007; Harberle et al., 2008
Trifolium	X	PS	Present	Present	Cai et al., 2008

[a] X, gene seems to be absent from the chloroplast genome; PS, a pseudogene is present in the chloroplast genome; Present, a fully functional gene is found in the chloroplast genome.

[b] *clpP1* in *Cynodon* is reported to include all of the coding region, but is missing both introns.

class inhibit translation of the chloroplast genome by preventing peptide bond formation (Douthwaite, 1992). Lincosamides function by binding to the 23S rRNA in the 50S ribosomal subunit, and disassociating peptidyl-tRNAs from the ribosome (Menninger and Coleman, 1993; Tenson et al., 2003). Resistance to lincomycin has been found by mutating a specific adenine in the 23S rRNA (Douthwaite, 1992; Tenson et al., 2003). Streptomycin inhibits translation of the chloroplast genome by interfering with the initial selection and proof-reading of the aminoacyl-tRNA in the A-site of the ribosome (Carter et al., 2000; Wirmer and Westhof, 2006). Streptomycin binds to the sugar-phosphate backbone of the 16S rRNA in four locations, helices 1, 18, 27 and 44, and the S12 ribosomal protein (Wirmer and Westhof, 2006), and functions by stabilizing the A-site in the 30S subunit in a conformation that increases the affinity of binding of any aminoacyl-tRNA (Carter et al., 2000; Peske et al., 2004). Resistance has been found through mutations at multiple positions on the 16S rRNA including the 530 loop in helix 18, and the region around nucleotide 912 (Wirmer and Westhof, 2006). Mutations in the S12 ribosomal protein have also shown resistance, including a lysine residue that interacts with helix 44, and multiple amino acids within the loops that interact with regions on the 16S rRNA (Carter et al., 2000).

Tetracyclines are a group of antibiotics that inhibit chloroplast translation by blocking the binding of aminoacyl-tRNAs to the A-site of the ribosome (Brodersen et al., 2000; Wirmer and Westhof, 2006). Similar to spectinomycin, tetracycline binds to the minor groove of helix 34 of the 16S rRNA in the 30S ribosomal subunit along with helix 31 (Wirmer and Westhof, 2006). Rather than interfering with translocation like spectinomycin, the binding of tetracycline inhibits interaction of aminoacyl-tRNAs with the A-site (Brodersen et al., 2000; Wirmer and Westhof, 2006).

Pactamycin inhibits chloroplast translation by preventing the formation of the initiation complex in translation (Brodersen et al., 2000; Dinos et al., 2004; Wirmer and Westhof, 2006).

Specifically, pactamycin binds to the 16S rRNA at helices 23 and 24, and the S7 ribosomal protein, which causes the two helices to lock together and the mRNA in the E-site of the ribosome to be moved by 12 Å. This displacement of the mRNA is believed to block translocation of peptidyl-tRNAs into the E-site (Brodersen et al., 2000; Dinos et al., 2004; Wirmer and Westhof, 2006). Resistance to pactamycin has been found in mutations at positions A694, C795, and C796 in the 16S rRNA of *Halobacterium halobium* (Mankin, 1997; Wirmer and Westhof, 2006).

For this project, we chose spectinomycin to use in most of the experiments because it is the most widely used agent to inhibit chloroplast translation, especially in relation to chloroplast transformation. Lincomycin was chosen to confirm the results because it disrupts chloroplast translation by binding to the 23S rRNA rather than the 16S rRNA.

Structure and Function of Acetyl-CoA Carboxylases (ACCases)

Biotin-dependent carboxylases are a large class of enzymes that utilize a molecule of biotin to catalyze the transfer of CO_2 between substrates. Among this class of enzymes are ACCases, which function to convert acetyl-CoA to malonyl-CoA during fatty acid biosynthesis (Tong, 2013). Most plant species contain two different versions of ACCases that function in different steps in fatty acid biosynthesis. The first version, known as ACC1 in Arabidopsis, is a large, homomeric protein localized to the cytosol that functions in the formation of very long-chain fatty acids (VLCFA), which are used in the formation of cuticular waxes, seed storage compounds such as triacylglycerides, suberin and sphingolipids, flavonoids (Amid et al., 2012), and other secondary metabolic compounds (Baud et al., 2004; Lü et al., 2011; Amid et al., 2012; Shang et al., 2016). VLCFAs and their derivatives have also been found to play a role in signaling within plants to regulate programmed cell death (Raffaele et al., 2008), activate ethylene

biosynthesis to promote cell elongation (Qin et al., 2007), suppress cell proliferation in the epidermis (Nobusawa et al., 2013), and regulate callus formation in culture (Shang et al., 2016). Null mutations in *ACC1* in Arabidopsis result in embryo lethality with the embryos arresting as "green blimps" without a defined hypocotyl or cotyledons (Meinke, 1985; Baud et al., 2003). Weak mutations result in decreased cuticular wax, reduced fertility, glossy inflorescence stems, and cold sensitivity (Lü et al., 2011; Amid et al., 2012).

The second type of ACCase found in most plant species is a heteromeric protein, similar to ACCases in bacteria, which is localized to chloroplasts and functions in a critical, early step of *de novo* fatty acid biosynthesis (Tong, 2013; Salie and Thelen, 2016). In vascular plants, excluding grasses that utilize a homomeric, chloroplast-localized ACCase, the four functional domains of the heteromeric ACCase are encoded by individual genes. The biotin carboxylase (BC) domain, biotin carboxyl carrier protein (BCCP domain), and carboxyltransferase (CT) α domain are encoded by the nuclear genome, while the CT-β domain is encoded by the *accD* gene within the chloroplast genome (Gu et al., 2011; Li et al., 2011). Null mutations in the *CAC1A* gene in Arabidopsis, which encodes one isoform of BCCP, result in embryo lethality with the embryos arresting at early stages of development. This phenotype is not seen with null mutants of *CAC1B*, a paralog to *CAC1A* (Li et al., 2011).

Some species of higher plants contain a homomeric version of ACCase that is localized to the chloroplast. Grasses have lost the heteromeric ACCase during the evolution of the Poaceae family, and contain only a chloroplast-localized, homomeric ACCase encoded by the nuclear genome (Jansen et al., 2007; Chalupska et al., 2008). This homomeric protein is the target for three classes of herbicides: aryloxyphenoxypropionates (FOPs), cyclohexanediones (DIMs), and phenylpyrazolins (DENs) (Kaundun, 2014). All three herbicide classes bind to the dimer interface within the CT domains of the protein and interfere with binding of acetyl-CoA (Zhang et al., 2004; Tong, 2013; Kaundun, 2014). Resistance to these herbicides has been found in plants

containing mutations in the two CT domains; specifically positions 1781, 2027, 2041, 2078, and 2096 (Liu et al., 2007). Most members of the Brassicaceae family have a duplicated copy of the homomeric, cytosolic ACCase that is targeted to chloroplasts. This means that Brassicaceae species contain three functional ACCases: one homomeric, cytosolic protein; one homomeric, chloroplast-localized protein; and one heteromeric, chloroplast-localized protein (Babiychuk et al., 2011; Bryant et al., 2011; Parker et al., 2014). Chloroplast-localized, homomeric ACCases can also be found in some algal species in the Prasinophyceae group that is thought to have been acquired through horizontal gene transfer rather than gene duplication (Huerlimann et al., 2015).

As noted above, ACCase proteins are composed of four main domains: BC, BCCP, CT-α, and CT-β (Tong, 2013). The BC domain catalyzes the first step in the conversion of acetyl-CoA to malonyl-CoA through ATP-dependent carboxylation of a biotin molecule, which is covalently bound to a specific lysine residue in the BCCP domain (Ohlrogge and Browse, 1995; Tong, 2013; Zu et al., 2013). This is shown at the top of Figure 3, where biotin is bonded to the BCCP domain, and is shown in the active site of the BC domain receiving a carboxyl group. There are three sub-domains within the BC domain (A, B, and C). The active site for the carboxylation step is located in the A and C sub-domains while the B sub-domain acts as a lid, and folds over the active site during the carboxylation (Tong, 2013; Zu et al., 2013). The function of BCCP region is to covalently bind the biotin molecule through biotinylation, and allow for translocation of this molecule between the BC and CT domains (Tong, 2013; Zu et al., 2013). The two CT domains work together to catalyze the second step in the carboxylation of acetyl-CoA to form malonyl-CoA through the transfer of the activated carboxyl group from the carboxybiotin molecule to a molecule of acetyl-CoA resulting in the production of malonyl-CoA (Ohlrogge and Browse, 1995; Tong, 2013; Zu et al., 2013; Wei and Tong, 2015). This step is shown at the bottom of Figure 3. The biotin molecule with its added carboxyl group is translocated to the active site within the pocket created by the CT domains, and the carboxyl group is transferred from biotin to

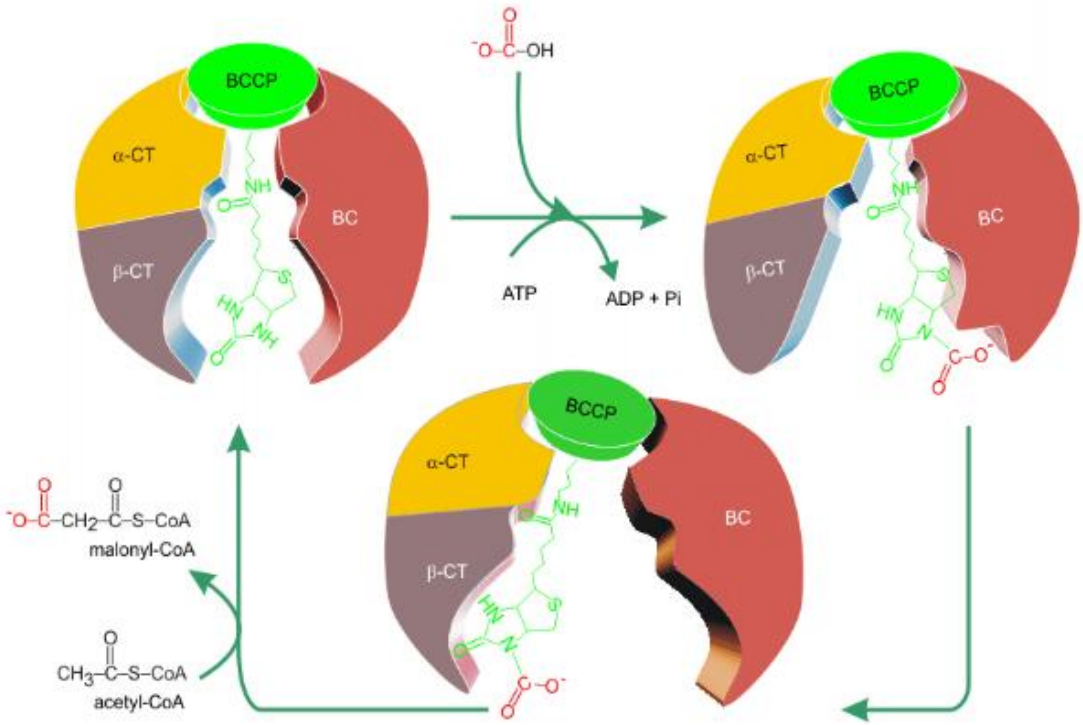


Figure 3. Biochemical Conversion of Acetyl-CoA to Malonyl-CoA Driven by an ACCase Enzyme. This image shows the cyclical process of forming malonyl-CoA from an activated carboxyl group attached to a biotin molecule and a free acetyl-CoA molecule. Adapted from the Acetyl-CoA Carboxylase webpage at The Arabidopsis Acyl-Lipid Metabolism Website (http://aralip.plantbiology.msu.edu/hehos/2).

an acetyl-CoA molecule, which becomes malonyl-CoA. The active site for this transfer of the activated carboxyl group is located inside an opening created through dimerization of the CT domains (Tong, 2013; Zu et al., 2013).

Homomeric ACCase proteins function as a dimer, which is essential for the catalytic reactions of the BC and CT domains (Figure 4). Monomers of the eukaryotic BC domain, tested *in vitro*, showed zero catalytic activity even though the monomers still had a high affinity for binding soraphen, a molecule that inhibits function of the BC domain (Weatherly et la., 2004; Wei and Tong, 2015). Large conformational changes have been found between the structures of the BC domain in the monomer and the BC domains in the dimer. This difference is believed to explain the inactivity of the monomer (Weatherly et la., 2004; Tong, 2013; Wei and Tong, 2015). The active site of the CT domain is formed by dimerization, which creates an opening surrounded by two CT-α and two CT-β domains (Bilder et al., 2006; Tong, 2013; Zu et al., 2013).

Chloroplast Protein Import via the TIC/TOC Import System

Thousands of proteins encoded by the nuclear genome function in the chloroplast. These proteins are imported into the chloroplast through the TIC/TOC (Translocon at Outer envelope membrane of the Chloroplast) protein import system (Shi and Theg, 2013). The TIC and TOC complexes are composed of numerous membrane-bound proteins, and chaperone proteins that work with them (Jarvis, 2008; Kessler and Schnell, 2009; Li and Teng, 2013; Shi and Theg, 2013). Both complexes exist in at least two different forms, with some redundancy between them: one that imports primarily housekeeping proteins into the chloroplast and one that imports photosynthetic proteins (Constan et al., 2004; Inoue et al., 2010; Hirabayashi et al., 2011; Kasmati et al., 2011). Housekeeping proteins appear to be imported mainly through complexes that include Toc34, Toc132/ Toc120, and Tic20-IV whereas photosynthetic protein import

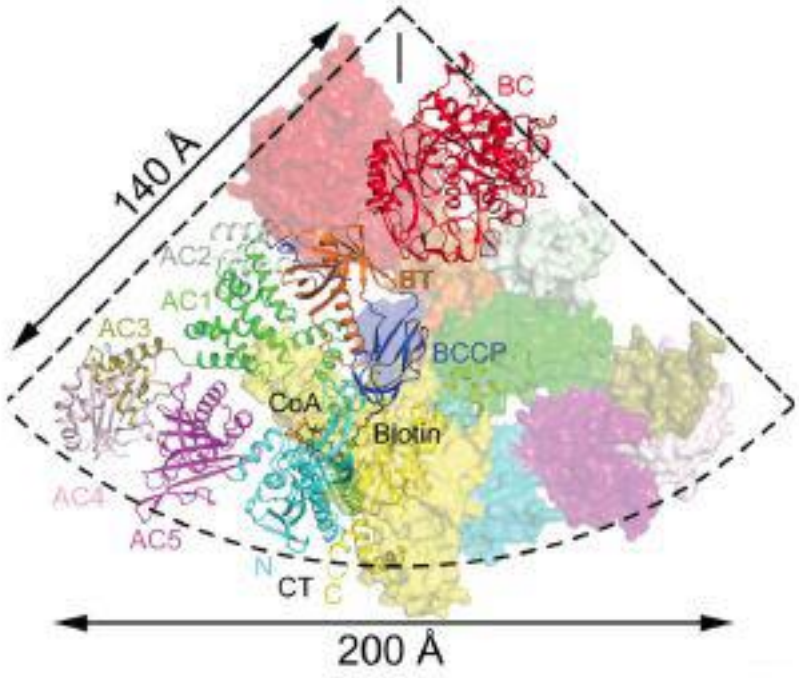


Figure 4. Crystal Structure of a Yeast ACCase Dimer. This image shows the crystal structure of a Yeast ACCase holoenzyme dimer. The two monomers are shown separately; one as a ribbon structure and one as a surface structure. The colors within the image correspond to the different domains: Red, BC; Blue, BCCP; Turquoise, β-CT; Yellow, α-CT; and Greens/PinkPurple, central domain. Adapted from Wei and Tong (2015).

involves Toc33, Toc159, and Tic20-I (Hirabayashi et al., 2011; Kasmati et al., 2011). Understanding this system is important for this project since the import of ACC2 into the chloroplast is needed to compensate for a loss of chloroplast translation. Components of the TIC/TOC protein import system are also potential modifiers that enhance tolerance of a loss of chloroplast translation. As the ACC2 precursor protein flows through the TOC and TIC import complexes of the chloroplast, there are a number of different proteins, detailed below, that likely interact with ACC2. These interactions can affect how the protein is folded, the stability of the precursor protein as it passes through the chloroplast membranes, and how the protein is recognized for translocation into the chloroplast.

The TOC protein complexes are composed of a protein conducting channel (Toc75), and two receptor GTPases (Li and Teng, 2013; Shi and Theg, 2013). Toc75 seems to function in both versions of the TOC protein complex as a guide for proteins going through the outer membrane of the chloroplast (Huang et al., 2011). Two isoforms of Toc75 are found in the Arabidopsis genome. The main isoform, encoded by *AtTOC75-III*, functions as the protein conducting channel, Toc75, while the second version, Outer Envelope Protein 80 kDa (OEP80) encoded by *AtOEP80*/*AtTOC75-V*, and has an unknown function (Huang et al., 2011; Shi and Theg, 2013). Null mutations in *AtTOC75-III* and *AtOEP80* result in embryo lethality at early stages of development, meaning that both proteins are required for embryo development in Arabidopsis (Baldwin et al., 2005; Patel et al., 2008; Meinke et al., 2009).

Toc159 and Toc33 are two GTPases found in the TOC protein complex typically associated with the import of photosynthetic proteins. Their counterparts associated with the import of housekeeping proteins are Toc132/ Toc120 and Toc34, respectively (Hirabayashi et al., 2011; Shi and Theg, 2013). Toc159 and Toc132/ Toc120 are composed of three domains: (1) the M domain, the C-terminus that anchors the protein to the outer membrane; (2) the G domain, the GTP-binding location; and (3) the A domain, the N-terminus (Kubis et al., 2004). Inoue et al.

(2010) have shown that the A domain on Toc159 and Toc132 heavily influences the selectivity of the TOC protein complexes for either photosynthetic or housekeeping proteins. A fourth member of the Toc159-type GTPases, Toc90, lacks an A domain and has been shown to function similarly to Toc159 at low levels (Hiltbrunner et al., 2004; Infanger et al., 2011). Overexpression of Toc90 can partially rescue *ppi2*, a knockout of Toc159 (Infanger et al., 2011). Similar to Toc159 and Toc132/Toc120, Toc34 and Toc33 primarily function as receptors for precursor proteins (Kubis et al., 2003; Constan et al., 2004; Shi and Theg, 2013). Members of the TOC import complex are some of the first proteins to interact with ACC2 as it is being translocated into the chloroplast. Recognition of ACC2 by these proteins is crucial for the translocation to occur (Kubis et al., 2004; Inoue et al., 2010).

The TIC protein complexes are composed of several proteins that work together to form the protein conducting channel across the inner chloroplast membrane (Hirabayashi et al., 2011; Kikuchi et al., 2013; Li and Teng, 2013; Shi and Theg, 2013). Tic20 is believed to be one of the channel proteins for translocation across the inner membrane, and has been shown to form a 1-megadalton (MDa) complex with Tic56, Tic100, Tic214, and potentially Tic21 (Kasmati et al., 2011; Kikuchi et al., 2013).

Four genes encoding different isoforms of Tic20 can be found in Arabidopsis. Little is known about the function of the proteins encoded by of two of these genes, *AtTIC20-II* and *AtTIC20-V*, which are expressed at high levels throughout plant development (Kasmati et al., 2011; Shi and Theg, 2013). The other two genes, *AtTIC20-I* and *AtTIC20-IV*, are thought to play crucial roles as channels for the import of photosynthetic and housekeeping proteins, respectively (Hirabayashi et al., 2011; Kasmati et al., 2011; Kikuchi et al., 2013). Not much is known about Tic21, but there is evidence it plays a role in the assembly of the 1-MDa complex (Teng et al., 2006; Shi and Theg, 2013). There is a debate over a second function of Tic21 in iron transport across the chloroplast membrane (Shi and Theg, 2013). This hypothesis was introduced by Duy et

al. (2007) when they showed that iron homeostasis-related proteins are upregulated in Arabidopsis *tic21* mutants. On the other hand, Kikuchi et al. (2009) maintain that Tic21 functions solely in the TIC protein import complex since the upregulation is also found in *tic20* and *albino3* mutants.

Tic110, Tic40, and the stromal chaperone protein Hsp93 are thought to function together as the translocation motor in the stroma (Kovacheva et al., 2005; Shi and Theg, 2013). Null mutations in *AtTIC110* result in embryo lethality at an early stage of development, which is consistent with the function of Tic110 as a recruiter for stromal chaperone proteins (Kovacheva et al., 2005). Tic40 is believed to be a chaperone to Tic110 where it binds to the protein in order to encourage the release of the transit peptide from the precursor protein being imported (Chou et al., 2006; Shi and Theg, 2013). The transit peptide is then cleaved by the stromal processing peptidase (SPP) before the final folding of the protein (Trösch and Jarvis, 2011; Shi and Theg, 2013). Stengel et al. (2009) showed that further regulation of the protein import complexes is provided by Tic62, Tic55, and Tic32 through redox signaling derived from photosynthesis. Members of the TIC import complex are needed to finish translocating the ACC2 precursor protein into the stroma of the chloroplast so that it can be folded into the final ACC2 protein, and function in fatty acid biosynthesis.

There is evidence of numerous chaperone proteins that function throughout the TIC/TOC protein import system. Cytosolic chaperones, such as Heat-shock protein 70 kDa (Hsp70), are thought to assist with the movement of precursor proteins from the ribosome to the TOC import complex on the surface of the chloroplast (Flores-Pérez and Jarvis, 2013). Hsp70 has also been shown to be involved in degradation of targeted precursor proteins (Lee et al., 2009; Flores-Pérez and Jarvis, 2013). Cytosolic chaperone protein 14-3-3 seems to complex with Hsp70 to help guide some types of precursor proteins to Toc34, which increases the efficiency of protein import for these proteins (May and Soll, 2000; Flores-Pérez and Jarvis, 2013). Hsp90 and AnKyrin

Repeat-containing protein 2 (AKR2) are additional cytosolic chaperones that are believed to function in guiding precursor proteins to the TOC import complexes (Flores-Pérez and Jarvis, 2013). Tic22 is thought to chaperone precursor proteins across the intermembrane space (IMS) between the TOC and TIC import complexes; although not much is known about the mechanism of this translocation (Kouranov et al., 1998; Shi and Theg, 2013). Stromal chaperones Hsp93, cpHsp70, and Hsp90C have been shown to operate alongside Tic110 and Tic40, and provide the driving force to translocate precursor proteins into the stroma (Kovacheva et al., 2007; Inoue et al., 2013; Shi and Theg, 2013). Translocation of ACC2 into the chloroplast cannot happen without chaperone proteins. These proteins are there to fold, stabilize, and guide ACC2 as it moves across the chloroplast membranes.

CHAPTER III

NATURAL VARIATION IN SEEDLING RESPONSES TO A LOSS OF CHLOROPLAST TRANSLATION IN ARABIDOPSIS

INTRODUCTION

Since the mid-1990s, natural variation among Arabidopsis accessions has been used to study a variety of fundamental questions in plant biology (Alonso-Blanco et al., 2009; Weigel, 2012). Various tools and resources are available for natural variation studies in Arabidopsis, including more than 7000 accessions available through seed stock centers, and whole-genome sequences for over 850 accessions (http://signal.salk.edu/atg1001/3.0/gebrowser.php; Weigel, 2012). In order to understand why plant species differ in their responses to a loss of chloroplast translation, we first looked to see if the phenotypic variation seen between Arabidopsis, Brassica, and tobacco could be found among natural accessions of Arabidopsis. We conducted two forward genetic screens analyzing seedling responses on spectinomycin, an inhibitor of chloroplast translation. Our original analysis of 52 accessions (Parker et al., 2014) was later expanded to include an additional 100 accessions chosen from the 1001 Genomes Project (Parker et al., 2016).

Selected accessions were also tested on lincomycin, a second antibiotic that inhibits chloroplast translation with a different mode of action, to confirm that the phenotypes seen on spectinomycin were caused by a loss of chloroplast translation. To further study the nuclear genes underlying tolerance of a loss of chloroplast translation, crosses were performed between three accessions tolerant of spectinomycin and one sensitive accession. Most of the data presented in this chapter have been published (Parker et al., 2014; 2016), except for the spectinomycin details listed in Appendices A and B.

MATERIALS AND METHODS

Plant Material and Growth Conditions

Seeds for wild-type accessions of Arabidopsis analyzed on spectinomycin were obtained from the Arabidopsis Biological Resource Center (ABRC; https://abrc.osu.edu/) at Ohio State University (Parker et al., 2014; 2016). Names and stock numbers for the accessions are listed in Appendix A. Seeds for the "Nossen" accession were obtained from wild-type plants that segregated in mutant populations (*emb3126-1* and *emb3137-1*) grown in our laboratory (Parker et al., 2014).

Mature seeds were germinated on plates containing a nutrient-agar medium following the protocol described by Meinke et al. (2009). The basal germination medium used was composed of Murashige and Skoog salts, 3% (w/v) glucose, and 0.8% (w/v) agar. For growth on spectinomycin and lincomycin plates, 50 mg L^{-1} spectinomycin or 200 mg L^{-1} lincomycin was added to the autoclaved basal medium through sterile filtration immediately before pouring the plates (Parker et al., 2014). Prior to plating on the nutrient-agar plates, seeds were surface sterilized in 95% ethanol for 30 seconds followed by a treatment of 50% non-concentrated Clorox bleach (including 1 drop of Tween 20 detergent per 10 mL of bleach) for 6 minutes (Meinke et

al., 2009). Seeds were then washed several times with sterile water, and plated on round petri dishes (100 mm in diameter). For basal plates, 50 seeds were evenly spread across the plate, and for spectinomycin or lincomycin plates 20-30 seeds were evenly spread. Laminated templates were used to ensure seeds were in the same positions on each plate. After plating the seeds, the plates were stored at 4° C in a refrigerator for 2-3 days. For accessions that needed an extended germination period, the plates were stored in a refrigerator for 7 days. Once removed from the refrigerator, the plates were placed under fluorescent lights for 14-21 days at room temperature. Seedlings were then transplanted to pots containing a mixture of 12-parts vermiculite, 3-parts soil, and 1-part sand. Pots were placed under fluorescent lights set to 16h-light/8h-dark cycles in a growth room maintained at 23°C ± 1°C. Daily watering of the pots was done using a nutrient solution (0.35 g L^{-1}) of Excel 15-5-15 fertilizer (Scotts Miracle-Gro, Port Washington, NY, USA; Berg et al., 2005). Pots were partially submerged (to a depth of 0.5 to 1 inches) in nutrient solution and soaked for several minutes before draining. After 2-3 weeks in the growth room, plants requiring vernalization were transferred to a cold room for 5-6 weeks at 5°C under fluorescent lights set to 8h-light/16h-dark cycles. These plants were then returned to the growth room for flowering. For seed collection, dried siliques were typically harvested from individual plants. In some cases, bulk dry seeds were also harvested from groups of sibling plants. Seed stocks were stored in capped vials (2 mL Fisherbrand™ Free-Standing Microcentrifuge Tubes) in the refrigerator at 4°C.

Seedling Responses on Spectinomycin and Lincomycin

Responses of seedlings grown on antibiotics were evaluated 5 weeks after plating, with accommodation for plates refrigerated longer. Measurements were performed under a Wild (M7) dissecting microscope equipped with an ocular micrometer. Using a ranking system, the extent of

leaf development for each seedling was determined by the size and number of leaves produced. Six ranks were used to classify the seedlings: A, cotyledons only (no visible leaf initials); B, first pair of leaf initials (≤ 1.5 mm combined leaf span); C, multiple leaf initials (≤ 2.5 mm combined for the two largest initials including any callus growth); D, one pair of leaves (> 1.5 mm combined); E, multiple leaves (> 2.5 mm and ≤ 6 mm combined for the two largest); and F, multiple leaves (> 6 mm combined for the two largest). Leaf development was also measured by length (mm) and width (mm) of the largest developed leaf, and the number of leaves found in each category based on leaf length: A, < 1.5 mm; B, ≥ 1.5 mm and < 3 mm; C, ≥ 3 mm and < 4.5 mm; D, ≥ 4.5 mm and < 6 mm; and E, ≥ 6 mm. The leaf count was removed from later seedling screens because it was redundant information for the extent of seedling growth. Root development was measured by approximating the root length using 5 categories: A, < 2 mm; B, ≥ 2 mm and < 4 mm; C, ≥ 4 mm and < 6 mm; D, ≥ 6 mm and < 9 mm; and E, ≥ 9 mm. Observations were made for each seedling on the pigmentation of cotyledons and leaves, and the location of the root in the medium. On occasion, seedlings with evidence of slight greening were found, often caused by limited root contact with the medium. These seedlings were excluded from evaluation.

Seedling and Whole Plate Imaging

Seedling images were captured with a Nikon DXM1200 digital camera attached to a Wild M-8 dissecting microscope, using the Nikon ACT-1 version 2.51 software. Plates with lids removed were placed under the dissecting scope with a black background, and centered on the seedling imaged. Most images were captured at 12x magnification; 6x magnification was also used to capture the full extent of growth for larger seedlings. Whole plate images were taken with a Canon PowerShot SX30 IS digital camera attached to a copy stand equipped with tungsten

lights. Lids were removed, and the plates were placed on a black background. The background of published images was uniformly darkened to highlight the seedling using the GNU Image Manipulation Program (GIMP) version 2.8.2.

Crosses Between Different Wild-Type Accessions

Crosses between wild-type accessions were performed using a tolerant accession (Jl-3, Be-1, or Tsu-0) as the female, and a sensitive accession ("Nossen") as the male. Successful crosses were confirmed by PCR genotyping (Parker et al., 2014). Crosses were accomplished following the protocol described by Meinke et al. (2009). Late floral buds, with a developed ovary and non-dehiscent anthers on the female parent were carefully emasculated by removing the 6 anthers with fine-tipped (Inox No. 4) forceps under a Wild (M7) dissecting microscope. Pollen from the male parent was brushed across the stigma surface of the emasculated bud until the surface was covered. In order to identify the crossed silique after it matured, and to prevent pollen contamination, 1-4 open flowers immediately below the cross were removed from the stem. Lateral branches not containing a cross were removed from the female plant to direct nutrients to the branches containing crosses. Typically, 1-4 crosses were performed on a single plant, and there were 4-10 crosses within one pot of plants. After crossing, these pots were placed under fluorescent lights (16h-light/8h-dark cycle) in a Percival (Perry, IA USA) plant growth chamber (AR-36L) maintained at 22°C ± 1°C, and watered with the same nutrient solution used in the growth room. After 4-5 weeks in the growth chamber, dry siliques from the female plant, comprising the expected cross and the surrounding selfed siliques, were harvested and stored in the refrigerator at 4°C.

RESULTS

Arabidopsis Accessions Differ in Seedling Sensitivity to Spectinomycin

We chose the 52 Arabidopsis accessions for our original analysis based on several criteria: (1) short flowering time to simplify the analysis of genetic crosses; (2) broad geographic locations; (3) background accessions of mutants defective in chloroplast translation (Bryant et al., 2011); and (4) high genetic diversity based on previous studies of natural variation (McKhann et al., 2004; Nordborg et al., 2005; Clark et al., 2007). One of the accessions, derived from segregating populations of RIKEN insertion mutants, was designated "Nossen" because it differed from the sequenced Nossen accession, No-0 (Parker et al., 2014). Tolerance of accessions to spectinomycin was analyzed using a ranking system (A-F) to characterize the development of seedlings grown for five weeks on 50 mg/L spectinomycin and 30 g/L glucose (Figure 5).

Consistent with our expectation that natural accessions might differ in their ability to tolerate a loss of chloroplast translation, we found that these 52 accessions had a broad range of seedling phenotypes on spectinomycin (Table 2; Figure 6). Seedlings from the most tolerant accessions grew into albino rosettes containing multiple large leaves. At the other end of the spectrum, seedlings from the most sensitive accessions developed only rudimentary leaf initials or lacked such initials altogether. Between these two extremes, seedlings from intermediate accessions showed moderate leaf development. Examples of sensitive, intermediate, and tolerant seedlings can be seen in Figure 3. Even though we classified each accession as sensitive, intermediate or tolerant, the range of accessions was continuous from the most tolerant to the most sensitive. Within each accession, the seedling phenotypes were mostly consistent, except for some intermediate accessions that showed a broad range from sensitive to tolerant seedlings. This consistency within an accession is shown in Figure 7. Occasional seedlings with greening on

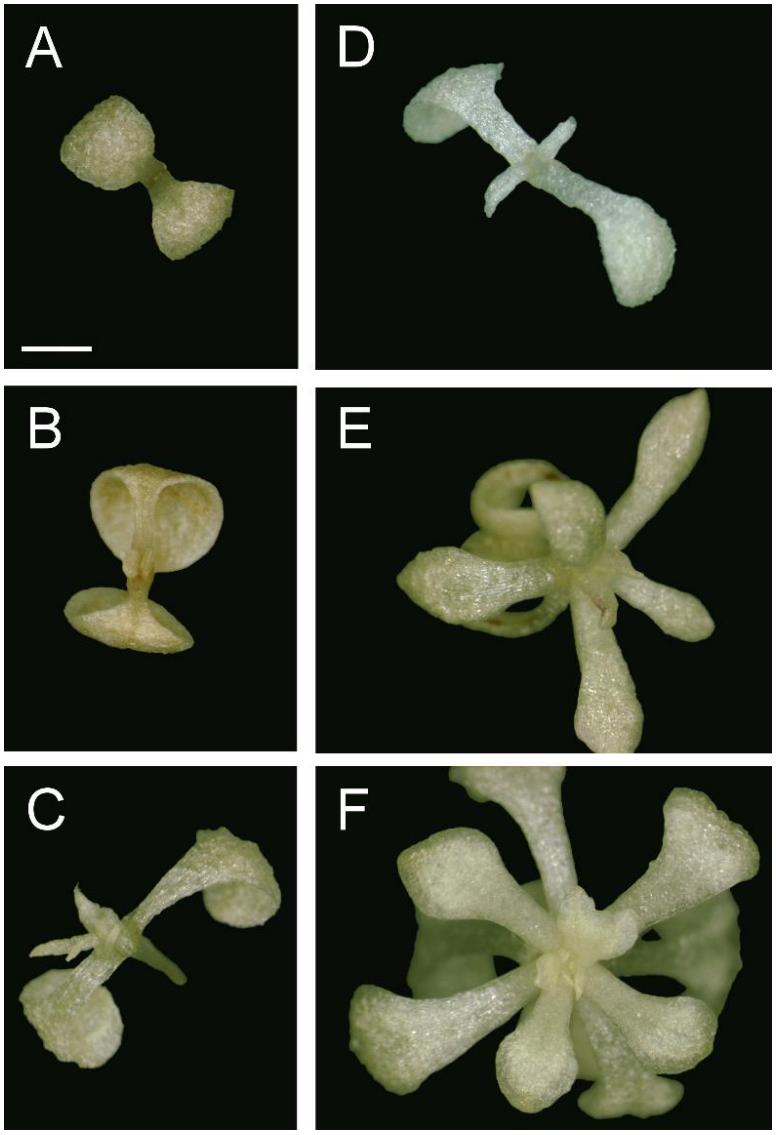


Figure 5. Seedling Phenotypes Reflecting Classification System. A and B, Sensitive accessions (categories A and B respectively). C and D, Intermediate accession (categories C and D respectively). E and F, Tolerant accessions (categories E and F respectively). Bar = 1 mm. Adapted from Parker et al. (2014; 2016).

Table 2. Seedling Responses of 52 Arabidopsis Natural Accessions Germinated on Spectinomycin. Additional details for all 52 accessions are presented in Appendix B. Adapted from Parker et al. (2014).

Accession Response Category	Total Accessions Classified	Total Seedlings Classified	Distribution of Seedling Phenotypes on Spectinomycin (%)[a]					
			Sensitive		Intermediate		Tolerant	
			A	B	C	D	E	F
Tolerant	20	613	1.3	1.6	3.8	2.0	**65.4**	**25.9**
Intermediate	15	409	5.4	13.2	**37.4**	15.2	**26.4**	2.4
Sensitive	17	526	**30.0**	**40.1**	12.5	15.0	2.3	

[a] Letters define classes from expanded cotyledons without leaves (A) to extensive rosettes with sizeable leaves (F) as defined in the text. Refer to Figure 3.1 for examples of seedling phenotypes for each class. Bold font, most common phenotypes (> 20%).

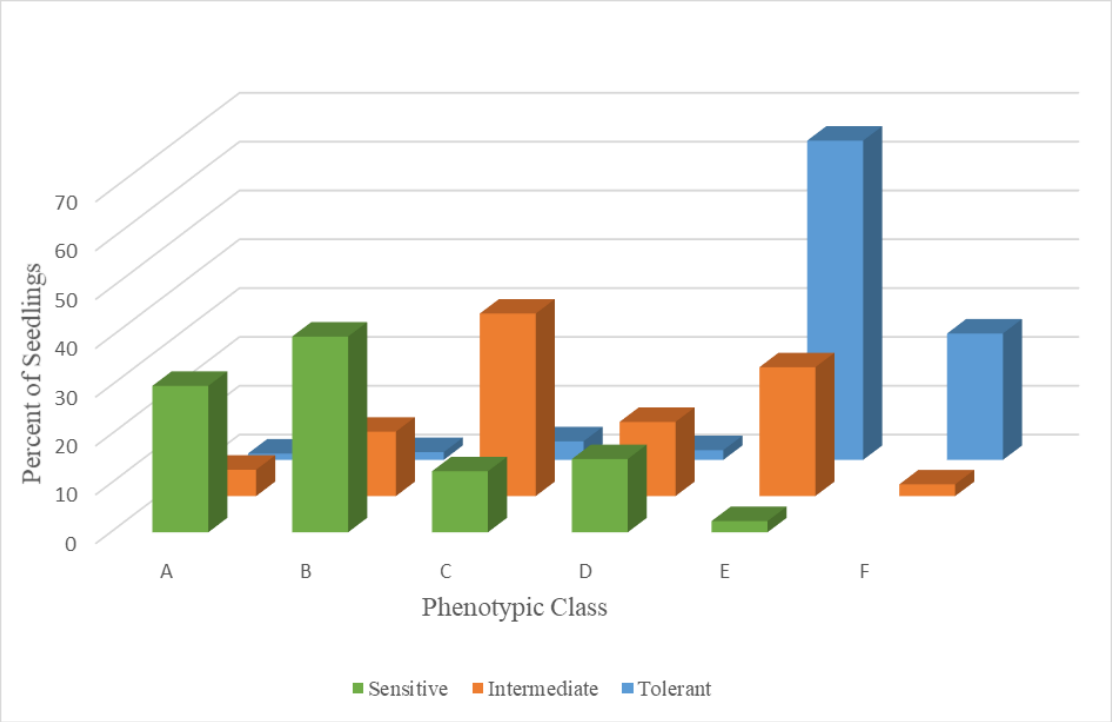


Figure 6. Spectinomycin Seedling Responses of 52 Arabidopsis Natural Accessions. Percent of seedlings in each class (Green, Sensitive; Orange, Intermediate; and Blue Tolerant) assigned to the six phenotypic categories (A-F) that are described in the Methods section of this Chapter. Additional data for these accessions can be found in Table 2 and Appendix B.

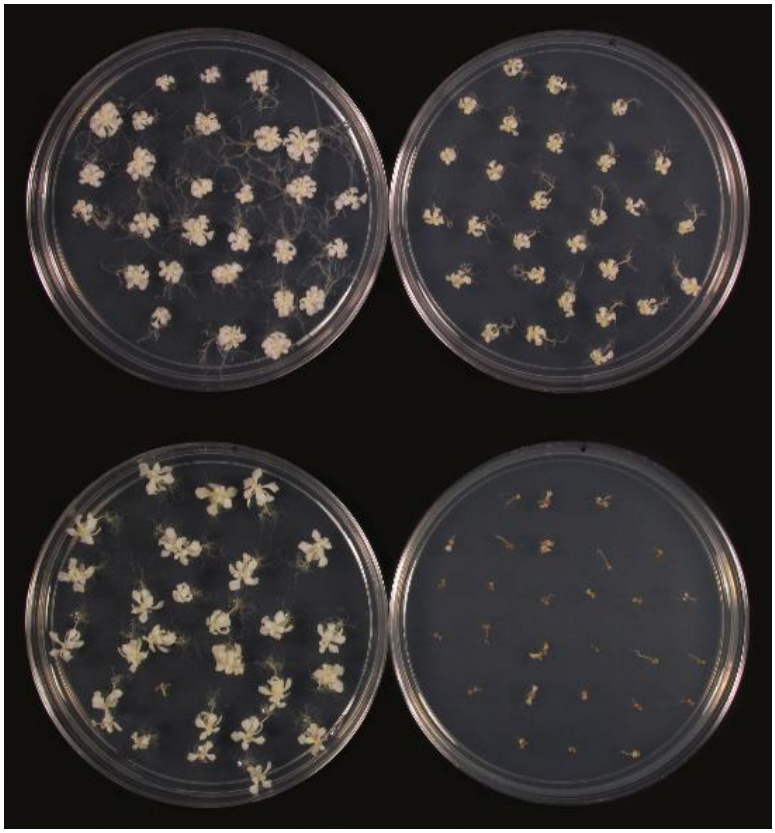


Figure 7. Consistent Seedling Responses of Arabidopsis Accessions on Spectinomycin. Clockwise from lower left: tolerant, Jl-3, Be-1, and Tsu-0; and sensitive, "Nossen". Plate diameter = 9 cm. Adapted from Parker et al. (2014).

cotyledons and leaves usually had poor root contact with the growth medium, which limited uptake of spectinomycin. With fewer than 20 seedlings tested for each accession in this original screen, minor growth differences between accessions were not significant. For further analyses, we concentrated on several of the most tolerant (Jl-3, Be-1, and Tsu-0) and sensitive ("Nossen", Oy-0, and Nie1-2) accessions identified. We initially excluded Sav-0 (the most sensitive accession) because genome sequence information was not available at that time. We also used Columbia (Col-0) because it is the most well-studied Arabidopsis accession, and it consistently shows an intermediate response. In order to confirm the range of tolerance found in accessions grown on spectinomycin, we tested these seven accessions on lincomycin, a second antibiotic with an entirely different mechanism to inhibit translation of the chloroplast genome. The extent of seedling growth for each accession on lincomycin mirrored the extent of growth on spectinomycin (Figures 8, 9). This supports our conclusion that differences in spectinomycin tolerance among natural accessions reflect fundamental differences in response to the inhibition of chloroplast translation.

To further study the nuclear genes underlying tolerance of a loss of chloroplast translation, we crossed wild-type plants from three tolerant accessions (Jl-3, Be-1, and Tsu-0) with the sensitive accession "Nossen". Progeny (F1) plants were allowed to self-pollinate, and the subsequent F2 seeds were plated on spectinomycin. Variation in the F2 seedling responses was observed for all three crosses examined (Figure 10). In all crosses, we could consistently identify sensitive seedlings that look like the "Nossen" parental, and tolerant seedlings similar to the tolerant parental. There was also a broad range of intermediate seedlings between the two phenotypes. This range in phenotypes was evidence of an underlying genetic basis for the phenotypic differences observed. Later, we focused solely on the cross between Tsu-0 and "Nossen" because these results were most similar to the 1:2:1 ratio expected for a single, semidominant genetic locus (Table 3; Figure 11).

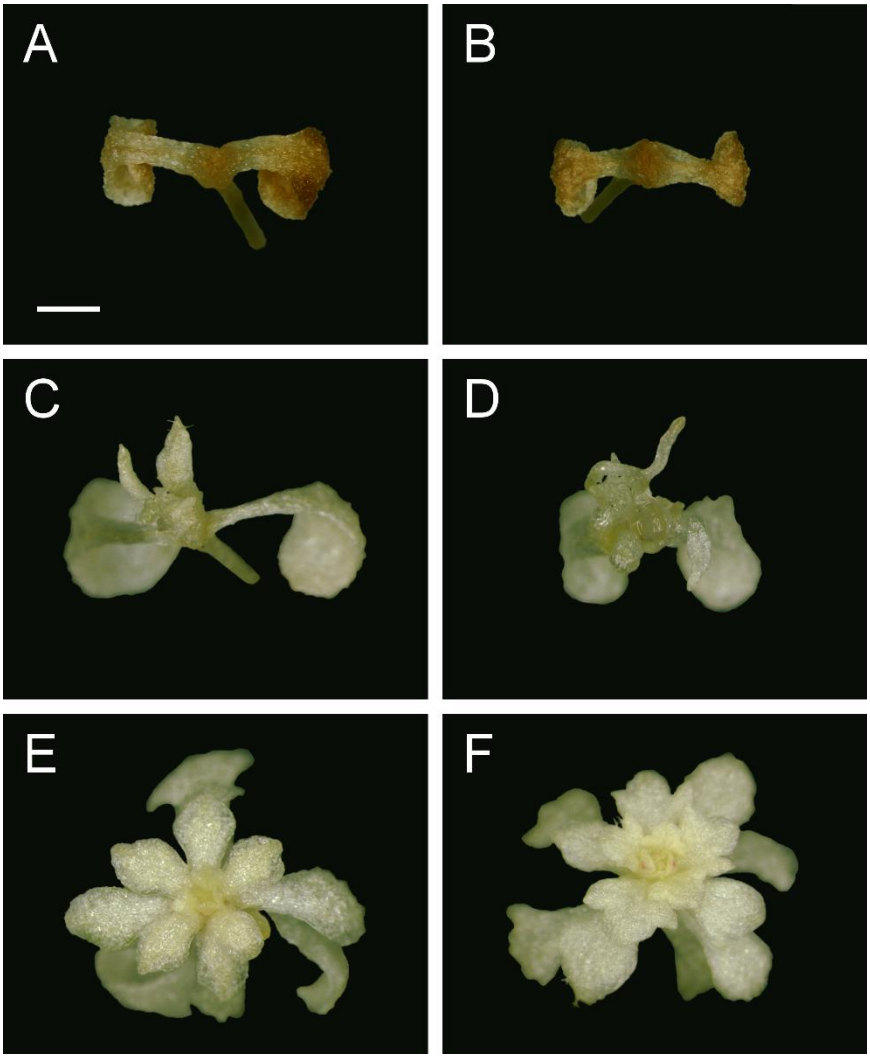


Figure 8. Seedling Responses of Three Arabidopsis Accessions on Spectinomycin and Lincomycin. A and B, sensitive accession, "Nossen", on spectinomycin (A) and lincomycin (B). C and D, intermediate accession, Col-0, on spectinomycin (C) and lincomycin (D). E and F, tolerant accession, Tsu-0, on spectinomycin (E) and lincomycin (F). Bar = 1 mm.

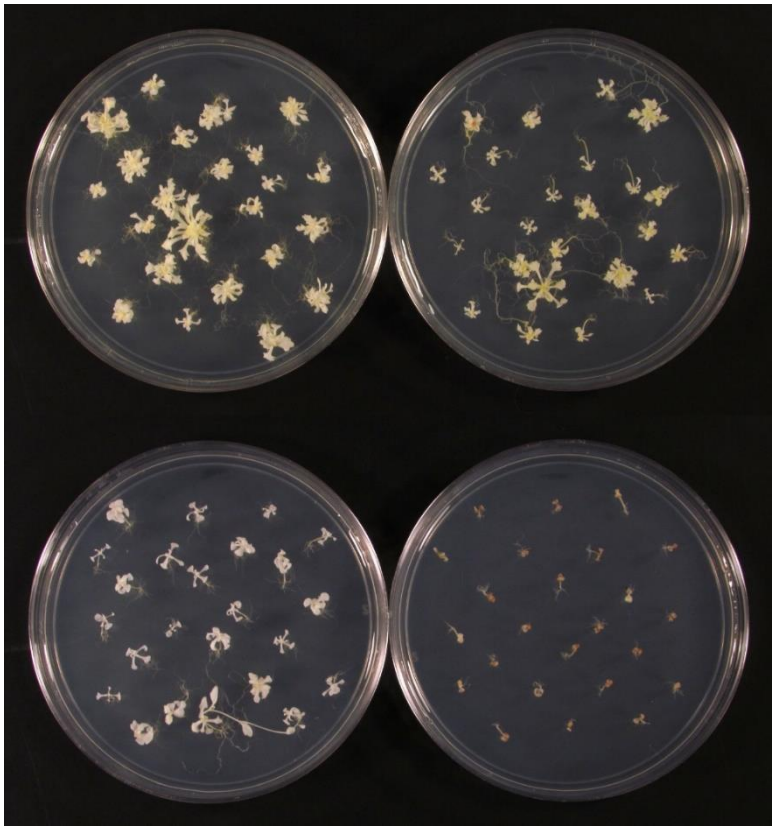


Figure 9. Consistent Seedling Responses of Arabidopsis Accessions on Lincomycin. Clockwise from lower left: tolerant, Jl-3, Be-1, and Tsu-0; and sensitive, "Nossen". Plate diameter = 9 cm. The consistency of response seen here is similar to that observed on spectinomycin (Figure 5). Adapted from Parker et al. (2014).

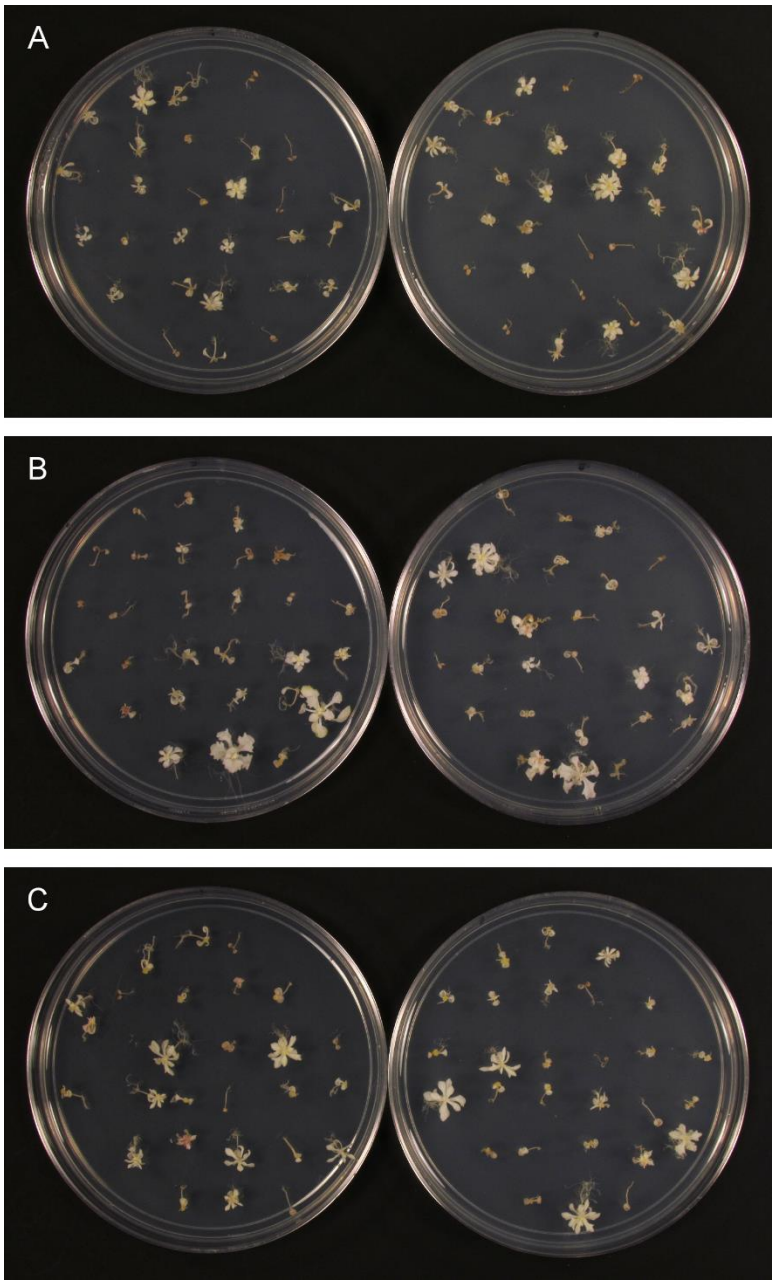


Figure 10. Segregating Seedling Responses in the F2 Generation from Crosses Between "Nossen" and Tolerant Accessions. A, “Nossen” crossed with Tsu-0. B, “Nossen” crossed with Be-1. C, “Nossen” crossed with Jl-3. Plate diameter = 9 cm. Adapted from Parker et al. (2014).

Table 3. Seedling Responses on Spectinomycin of Parental Accessions and F2 Progeny from Crosses Between "Nossen" and Tolerant Accessions. Adapted from Parker et al. (2014).

Genotype Examined	Total Seedlings Classified	Distribution of Seedling Phenotypes on Spectinomycin (%)[a]					
		Sensitive		Intermediate		Tolerant	
		A	B	C	D	E	F
"Nossen"	178	41.0	24.1	33.2	1.7		
Tsu-0	133					80.4	19.6
Tsu-0 x "Nossen"	233	23.2	0.9	21.5	3.4	42.0	9.0
Be-1	131					47.3	52.7
Be-1 x "Nossen"	140	15.0	12.9	23.6	2.8	37.9	7.8
Jl-3	135				0.7	31.9	67.4
Jl-3 x "Nossen"	198	23.7	8.6	15.7	11.6	28.3	12.1

[a] Letters define classes from expanded cotyledons without leaves (A) to extensive rosettes with sizeable leaves (F) as defined in the text.

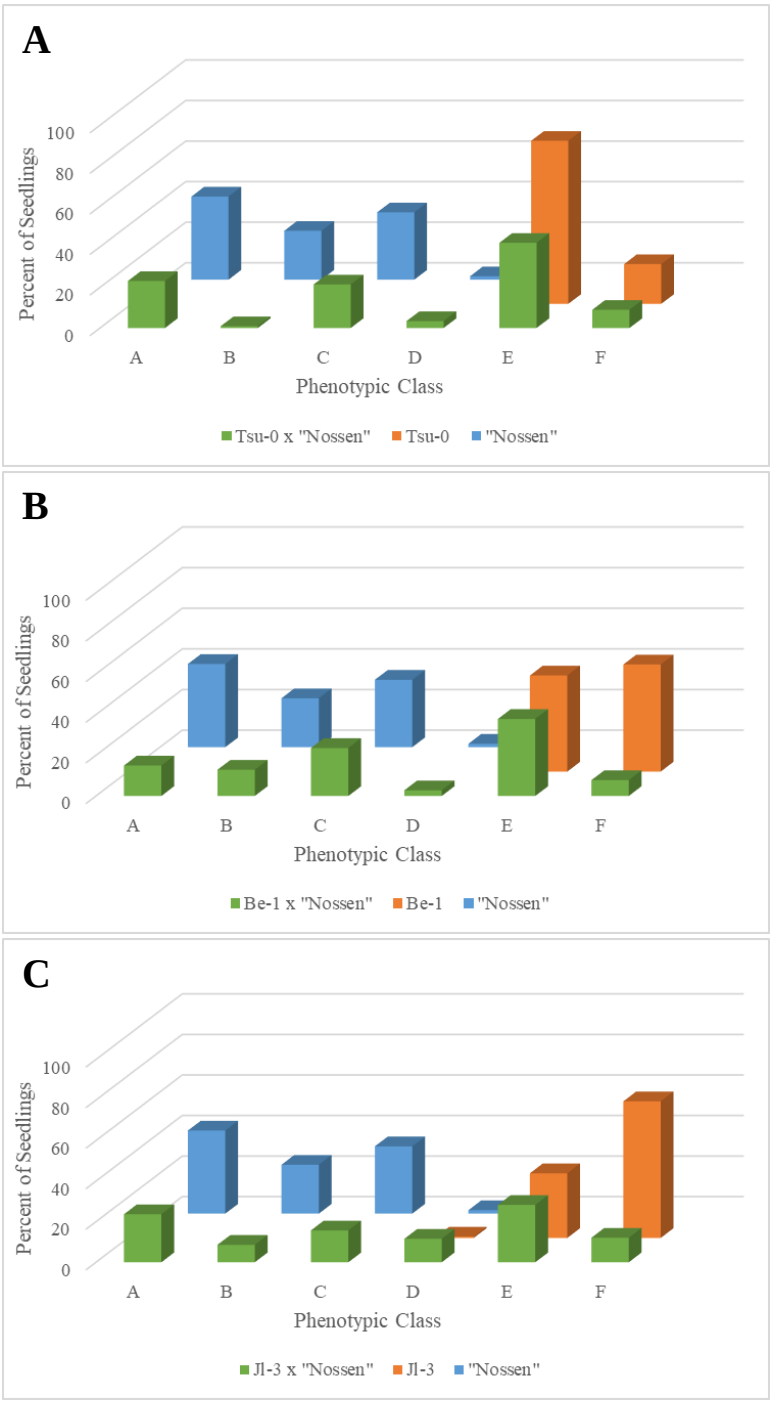


Figure 11. Comparison of Spectinomycin Seedling Responses of Parental Accessions and F2 Progeny from Crosses Between "Nossen" and Tolerant Accessions. Percent of seedlings in each accession or F2 line assigned to the six phenotypic categories (A-F) that are described in the Methods section of this chapter. A, Tsu-0 x "Nossen"; B, Be-1 x "Nossen"; and C, Jl-3 x "Nossen". The data for these crosses can be found in Table 3.

Evaluating Additional Lines Increases the Number of Sensitive Accessions

In order to learn more about what causes sensitivity of some accessions to a loss of chloroplast translation, we increased the total number of accessions tested on spectinomycin to identify additional sensitive accessions to study. For this second analysis, we chose 100 new accessions based on the following criteria: (1) availability of a sequenced genome from the 1001 Genomes Project; and (2) broad geographic locations (Figure 12). Seed stocks for these 100 accessions were derived from siblings of the plants sequenced in the 1001 Genomes Project.

The ranking system used to characterize the development of seedlings after five weeks of growth on spectinomycin was expanded from six categories to nine in order to create a quick method to calculate a phenotype score for each accession. Sensitive seedlings were classified as (1) cotyledons only (no visible leaf initials), (2) first pair of leaf initials (≤ 1.5 mm combined leaf span), or (3) multiple leaf initials (≤ 1.5 mm combined for the two largest initials including any callus growth). Intermediate seedlings were classified as (5) multiple leaves (> 1.5 mm and ≤ 2.5 mm combined for the two largest), (6) one pair of leaves (> 1.5 mm combined), or (7) multiple leaves (> 2.5 mm and ≤ 4 mm combined for the two largest). Tolerant seedlings, which all had multiple leaves, were classified as (9) > 4 mm and ≤ 6 mm, (10) > 6 mm and ≤ 9 mm, or (11) > 9 mm combined for the two largest. Examples of each seedling category can be seen in Figure 13. We calculated a phenotype score for each accession using the average rank of all individual seedlings measured. Utilizing the percentage of seedlings within each category, accessions were classified as hypersensitive (95% or more seedlings in categories 1 and 2 and 50% or more seedlings in category 1); sensitive (70% or more seedlings within a sensitive category); low intermediate (50% or more seedlings within a sensitive category); high intermediate (50% or more seedlings within a tolerant category); tolerant (70% or more seedlings within a tolerant category); or intermediate (everything that failed to meet any of the above criteria).

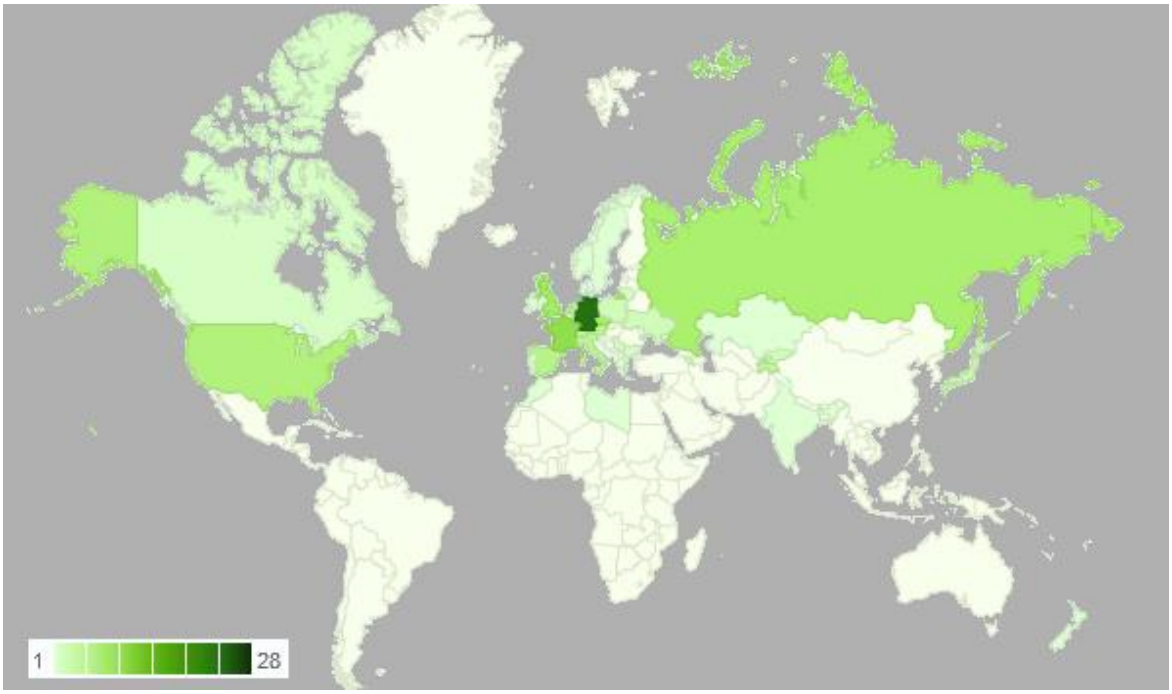


Figure 12. Global Distribution of 152 Natural Accessions Analyzed. White, no accessions used. Light Green, 1 or 2 accessions used. Darker green, 10-28 accessions used. Constructed using eSpatial Mapping Software (https://www.espatial.com/).

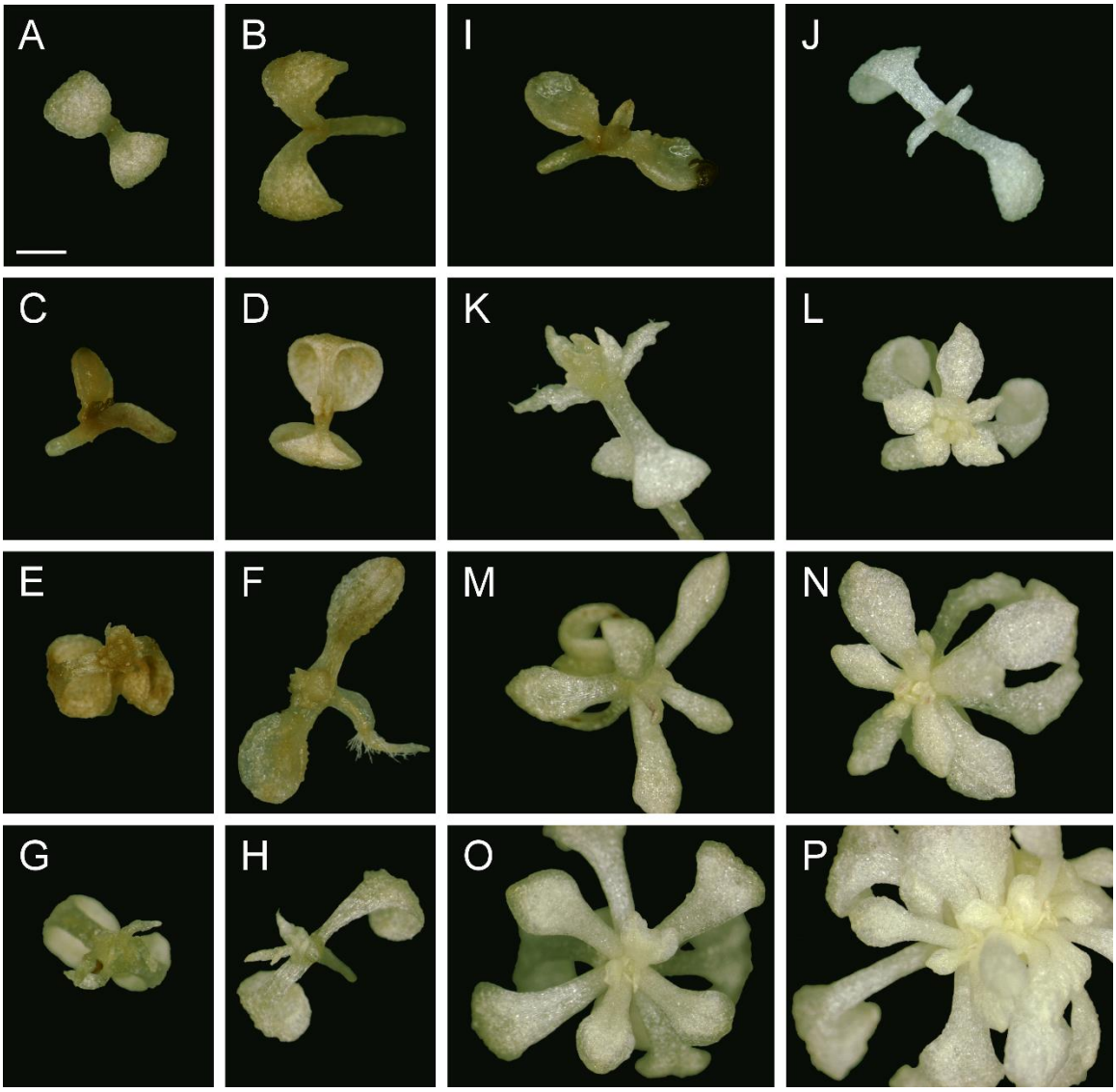


Figure 13. Seedling Responses of Selected Arabidopsis Accessions on Spectinomycin. A to F, Sensitive seedlings, categories 1 (A and B), 2 (C and D), and 3 (E and F). G to L, Intermediate seedlings, categories 5 (G and H), 6 (I and J), and 7 (K and L). M to P, Tolerant seedlings, categories 9 (M and N), 10 (O), and 11 (P). Bar = 1 mm. Adapted from Parker et al. (2016).

The 100 accessions from this second spectinomycin analysis showed the same broad range of seedling phenotypes as the first 52 accessions. Combining the results from both analyses, more the 8,000 seedlings from the 152 accessions were evaluated on spectinomycin (Table 4; Figure 14; Appendix B). Of these accessions, three were classified as hypersensitive, 22 as sensitive, 13 as low intermediate, 83 as intermediate, 11 as high intermediate, and 20 as tolerant. Again, consistency of seedling phenotypes was found within most accessions, except for some intermediate accessions that showed a broad range of seedling responses, for unknown reasons. Occasionally, a seedling from a tolerant accession grew poorly on spectinomycin showing a sensitive phenotype, possibly caused by poor nutrient uptake from the growth medium. On the other hand, hypersensitive accessions did not have any high intermediate or tolerant seedlings outside of those with greening and root problems, and sensitive accessions did not have any highly tolerant seedlings. Additional analyses, described later, focused on hypersensitive and sensitive accessions with the lowest phenotype scores recovered from a combination of forward and reverse genetic screens.

DISCUSSION

In this study, we used 152 natural accessions of Arabidopsis to explore the genetics underlying phenotypic differences found among plant species when translation of the chloroplast genome is blocked. Our results from spectinomycin studies of these 152 accessions show that differences originally reported by Zubko and Day (1998) between Arabidopsis, Brassica and tobacco, can also be found within Arabidopsis accessions. While a broad range of variation was found among the accessions, seedling phenotypes within an accession were mostly consistent, with the exception of some intermediate accessions that had a wide range. A number of these intermediate accessions possibly lack consistency due to the small number of seedlings analyzed

Table 4. Seedling Responses of 152 Arabidopsis Natural Accessions Germinated on Spectinomycin. Additional details for all 152 accessions are presented in Appendix B. Adapted from Parker et al. (2016).

Accession Response Category	Total Accessions Classified	Total Seedlings Classified	Accession Phenotype Scores	Distribution of Seedling Phenotypes on Spectinomycin (%)[a]								
				Sensitive			Intermediate			Tolerant		
				1	2	3	5	6	7	9	10	11
Tolerant	20	1,861	8.1 - 9.7	0.5	0.8	0.5	1.2	0.6	9.3	**52.0**	**30.0**	5.1
High Intermediate	11	477	6.4 - 8.3	2.5	1.0	0.4	4.6	1.7	**32.5**	**49.1**	8.0	0.2
Intermediate	83	2,824	3.9 - 7.8	5.2	8.0	6.4	**23.7**	9.8	**35.2**	**10.4**	1.3	
Low Intermediate	13	427	3.2 - 4.5	**12.9**	**33.7**	**12.2**	**15.2**	**15.9**	6.6	3.5		
Sensitive	22	1,872	1.3 - 3.2	**34.1**	**39.6**	**19.0**	3.5	2.4	1.0	0.4		
Hypersensitive	3	546	1.1 - 1.2	**86.8**	**10.6**	2.2		0.4				

[a] Numbers define classes from expanded cotyledons without leaves (1) to extensive rosettes with sizeable leaves (11) as defined in the text. Refer to Figure 3.7 for examples of seedling phenotypes for each class. Bold font, most common phenotypes (> 10%).

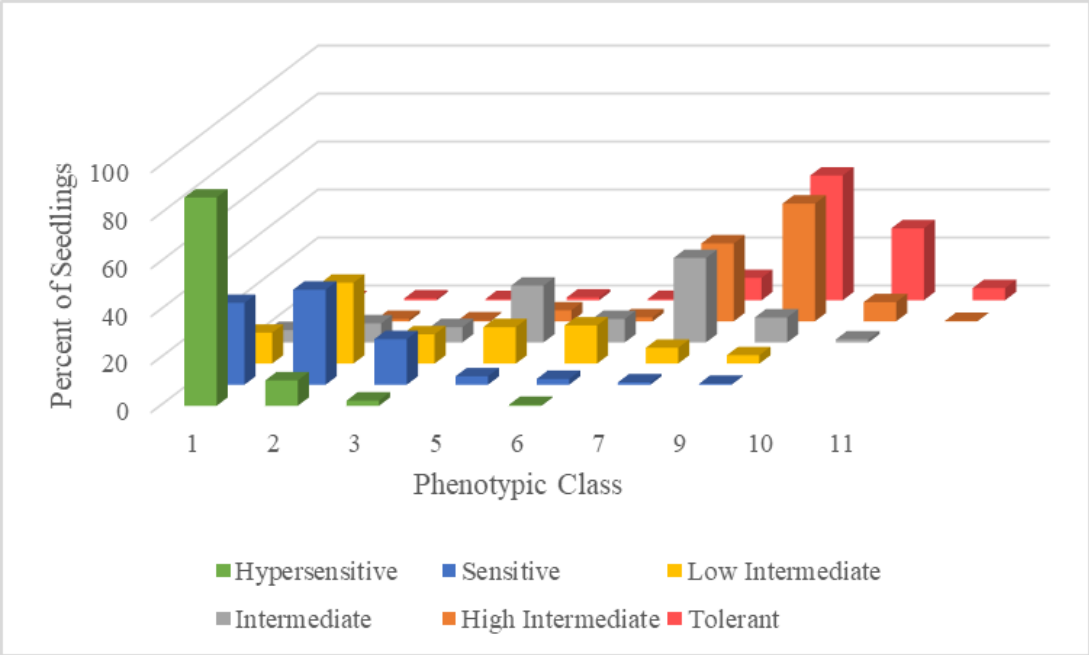


Figure 14. Spectinomycin Seedling Responses of 152 Arabidopsis Natural Accessions. Percent of seedlings in each class assigned to the nine phenotypic categories (1-3; 5-7; 9-11) that are described in "Evaluating Additional Lines Increases the Number of Sensitive Accessions" in this Chapter. Additional data for these accessions can be found in Table 4 and Appendix B.

per accession. Because this project focused on the most tolerant and sensitive accessions, many of the intermediates were not re-tested after the initial spectinomycin screen. For accessions that were evaluated more than once, there is the possibility that poor contact between the roots of the seedling and the spectinomycin media allowed for more extensive growth than was seen on other plates. Although, poor contact with the media typically resulted in greening of the seedling due to a decrease in spectinomycin uptake. Additional spectinomycin screenings and analyses of these intermediate accessions will be needed to determine if the lack of consistency in seedling phenotypes is due to plating inconsisitencies or something else.

The striking phenotypic differences observed here between the most tolerant and sensitive accessions provide a unique system to analyze the nuclear genes and cellular processes involved. We started to explore the genetic basis of these differences by crossing tolerant and sensitive accessions. The variation seen in the F2 seedling responses to spectinomycin from crosses between tolerant accessions (Jl-3, Be-1, and Tsu-0) and the sensitive "Nossen" accession suggests that there is a genetic component underlying the phenotype differences observed. While the crosses with Jl-3 and Be-1 were difficult to interpret, and seemed to indicate the involvement of multiple genes underlying the phenotype differences, the F2 responses from the cross between Tsu-0 and "Nossen" showed approximately a 1:2:1 ratio of tolerant to intermediate to sensitive seedlings, which would be expected if a single, semi-dominant genetic locus was underlying the phenotype differences between the parental accessions. However, there were limitations to using these crosses to identify the underlying gene(s). Since the F2 seedlings were analyzed on spectinomycin, we were not able to grow them in soil to harvest progeny (F3) seed. Seedlings at the borderlines between the sensitive, intermediate, and tolerant categories were hard to classify, which made it difficult to distinguish heterozygous seedlings, which should be intermediate, from those homozygous for either "Nossen" (sensitive seedlings) or Tsu-0 (tolerant seedlings). Since we could not readily use this system to identify specific genes, we turned to a different approach

that involved crossing the tolerant Tsu-0 accession with *emb* mutants (in the sensitive "Nossen" background) that were disrupted in chloroplast translation. This approach allowed us to harvest progeny of the same plants we were analyzing, and differences between plants were more distinct. The results of that approach are reported in Chapter 4.

Arabidopsis accessions hypersensitive to a loss of chloroplast translation clearly show that one or more genes in the chloroplast genome are essential for seedling development in Arabidopsis. We believe the most critical gene is *accD*, based on targeted gene disruptions in tobacco, and the retention of *accD* in the chloroplast genomes of parasitic plants. We later give further evidence that *accD* is the most critical gene in the chloroplast genome. Comparing the extent of development between tolerant accessions on spectinomycin and Arabidopsis mutants defective in photosynthesis (Bryant et al., 2011), the albino seedlings from tolerant accessions were not as extensively developed. This means that the loss of chloroplast translation in tolerant accessions is not fully rescued by *ACC2*. One possible explanation is that there are additional chloroplast gene(s) that become essential at later stages in seedling development. Among these candidate genes are *ycf1* and *ycf2*, which function in chloroplast protein import (Kikuchi et al., 2013; Parker et al., 2016). These genes might play a role in importing housekeeping proteins essential for later stages of plant development. Another candidate is *clpP1*, which is a subunit of a chloroplast-localized protease complex known to be required for chloroplast function (Ramos-Vega et al., 2015). All three of these genes, along with *accD*, were identified as essential chloroplast genes in tobacco (Drescher et al., 2000; Kuroda and Maliga, 2003; Kode et al., 2005).

CHAPTER IV

FACTORS THAT ENHANCE THE EXTENT OF EMBRYO DEVELOPMENT IN THE ABSENCE OF CHLOROPLAST TRANSLATION

INTRODUCTION

Following crosses between tolerant and sensitive wild-type accessions, we found that our procedure for screening F2 seedlings lacked the accuracy needed to identify the gene(s) responsible for phenotypic differences seen when chloroplast translation is blocked. Looking at 33 insertion mutants defective in both embryo development and chloroplast translation, we found a correlation between the stage of embryo arrest and the sensitivity of the parental accession when grown on spectinomycin (Table 5). Mutant embryos of RIKEN insertion mutants in a "Nossen" background arrest at a preglobular stage of embryo development, and wild-type "Nossen" seedlings are sensitive to a loss of chloroplast translation. On the other hand, mutant embryos of SALK or Syngenta insertion mutants in a Col-0 background arrest at a large globular stage of development, while wild-type Col-0 seedlings show an intermediate phenotype on spectinomycin. In between these two accessions, mutant embryos in CSHL or JIC insertion lines in a Ler-1 background arrest at a small globular stage of development, and wild-type Ler-1

Table 5. Chloroplast Translation Mutants Differ in Stage of Embryo Arrest. Adapted from Parker et al. (2014).

Accession	Insertion Line	Knockout Alleles	Embryo Phenotype	Embryo Mutants Used in Crosses	Size of Arrested Embryo	Ribosomal Protein
"Nossen"	Riken	6	Preglobular	*emb3126-1* *emb3137-1*	25 µm	L1 S13
Columbia	Salk/GABI	8	Large Globular	*emb3137-2*	90 µm	S13
Columbia	Syngenta	9	Large Globular	-	-	-
Ler	CSHL/JIC	2	Small Globular	*emb3126-3*	60 µm	L1

seedlings show a phenotype on spectinomycin that is in between sensitive "Nossen" seedlings and intermediate Col-0 seedlings. The correlation observed between the embryo and seedlings phenotypes when translation of the chloroplast genome is blocked is important because it suggests a common mechanism involved in both phenotypes.

Using this information, we were able to design a more accurate procedure to identify the gene(s) responsible for phenotypic differences seen when chloroplast translation is blocked. We performed crosses between wild-type plants of the tolerant Tsu-0 accession and plants segregating for an *emb* mutation that eliminated chloroplast translation in the sensitive "Nossen" accession. We focused on RIKEN insertion mutants in two *EMB* genes that encode chloroplast-localized ribosomal proteins, *EMB3126* and *EMB3137*, where the embryos arrest at a preglobular stage of development. Using these crosses, we screened for dominant suppressors of this preglobular arrest. Through this study, we found a single suppressor locus (*ACC2*), an enhancer of the suppressor, and additional modifiers that further increase embryo development. Most of the data presented in this chapter have been published (Parker et al., 2014). Two notable exceptions are the analysis of F5 embryos from crosses between Tsu-0 and *emb3126-1*, and details of the plants screened for mapping the enhancer locus and identifying additional modifiers (Appendices C, D).

MATERIALS AND METHODS

Plant Material

Details on the *emb* mutants used for this part of the project have been described in previous publications (Bryant et al., 2011; Muralla et al., 2011) and are presented in the SeedGenes database (http://www.seedgenes.org). Seeds for *emb3126-1* (RATM-53-3245-1), *emb3126-3* (GT-5-101962), *emb3137-1* (RATM-15-0663-1), and *emb3136* (RATM-51-2522-3)

were obtained from Kazuo Shinozaki at the RIKEN Plant Science Center. Seeds for *emb3137-2* (Salk-133412), *acc2-1* (Salk-148966c), and *acc2-2* (Salk-110264) were obtained from the ABRC (https://abrc.osu.edu/) at Ohio State University. Internal seed stocks were used for *emb1473* (Syngenta 24154) in the Columbia background; duplicates are available through the ABRC.

Crosses with *emb* Mutants and Embryo Phenotyping

Most of the crosses between wild-type accessions and plants heterozygous for an *emb* mutation (*emb/EMB*) were performed in both directions using the heterozygous *emb* plant as either the male or female. We identified heterozygous *emb* plants by screening mature siliques for the presence of 25% mutant seeds. When the heterozygous *emb* plant was used as the female parent, successful crosses were confirmed by the harvested silique lacking aborted seeds, which was different from the adjacent siliques produced from selfing. When the heterozygous *emb* plant was used as the male parent, successful crosses were determined by segregation of mutant F2 seeds in siliques of F1 plants. Seed and embryo measurements were taken under a Wild (M7) dissecting microscope using a stage micrometer and two fine-tipped (Dumont no. 4) forceps. The smallest embryos that we could measure this way were 50 μm globular embryos. Smaller embryos could be seen as bumps in the seed coat, but we were not able to dissect them out of the seed coat to measure. Mutant embryos were classified into four categories: (1) globular: rounded embryos; (2) triangular: embryos with a visible point at the basal region; (3) linear: embryos with elongation of the basal region without cotyledon formation; and (4) cotyledon: embryos with one or more cotyledons. In order to be sure that the embryos measured were at a terminal stage of development, we mostly dissected aborted seeds that had begun to deflate and turn brown.

Embryo Imaging

Embryo images were captured with a Nikon DXM1200 digital camera attached to a Wild M-8 dissecting microscope, using the Nikon ACT-1 version 2.51 software. Embryos were first extracted under a Wild (M7) dissecting microscope using two fine-tipped (Dumont no. 4) forceps, and placed on an open plate of medium to ensure the embryos did not dry out before imaging. Images were captured at 50x magnification. The background of published images was uniformly darkened to highlight the embryo using the GNU Image Manipulation Program (GIMP) version 2.8.2.

RESULTS

A Single, Dominant Suppressor of Preglobular Arrest Increases Seed and Embryo Development

In order to identify the nuclear genes that influence tolerance or sensitivity to loss of chloroplast translation, we focused on knockout mutants disrupted in two *EMB* genes required for chloroplast translation, *EMB3126* and *EMB3137*. These genes encode chloroplast-localized ribosomal proteins, L1 and S13 respectively, and both genes have mutant alleles defective in different genetic backgrounds with different embryo phenotypes (Table 6). We later discontinued the work on *emb3126-3*, which is in the Ler-1 background, due to the variable seed size in Ler-1. We crossed heterozygous (*emb/EMB*) plants from the RIKEN mutants, *emb3126-1* and *emb3137-1*, with the tolerant Tsu-0 accession, and screened for dominant suppressors of the preglobular arrest found in these mutants. A single dominant Tsu-0 suppressor should cause 75% of the mutant seeds in F1 siliques to reach a later stage of development. We expected to see three classes of segregating F2 plants: (1) those with an early seed phenotype similar to the *emb* parent; (2) those with a late seed phenotype; (3) and those with a mixture of both. If other modifiers were

Table 6. Mutant Alleles Chosen for Initial Crosses with Spectinomycin-Tolerant Accessions. Adapted from Parker et al. (2014).

Allele Symbol	Ribosomal Protein	Insertion Line	Background Accession	Embryo Phenotype	Embryo Size (μm)
emb3126-1	L1	Riken	"Nossen"	Preglobular	25
emb3126-3	L1	JIC	L*er*	Small globular	60
emb3137-1	S13	Riken	"Nossen"	Preglobular	25
emb3137-2	S13	Salk	Columbia	Large globular	90

involved, we expected to find F3 plants with more advanced embryos.

The crosses between Tsu-0 and *emb* mutants defective in chloroplast translation showed evidence of a single dominant suppressor that significantly increased the size of mutant seeds and supported embryo development to a late globular stage. The first two rows in Table 7 show the results of screening mutant seeds from F1 siliques. Around 75% of the mutant seeds screened from these crosses contained an embryo rescued to a large globular stage of development, while the other 25% were similar to the preglobular phenotype found in the parental *emb* mutant. When the next generation of plants was grown, three distinct classes of F2 plants were found: SS plants with a preglobular mutant seed phenotype similar to the *emb* parent; TT plants with a rescued (large globular or later development stage) seed phenotype; and ST plants with a mixture of rescued and parental seed phenotypes in a 3:1 ratio (Table 8). These F2 classes were found in a 1:2:1 ratio of SS:ST:TT plants. Because some embryo rescue was found in both the *emb3126-1* and *emb3137-1* crosses, the response is not limited to a specific ribosomal protein. As a control, we crossed *emb3126-1* and *emb3137-1* with two other tolerant accessions, Jl-3 and Be-1, and with two sensitive accessions, Oy-0 and Nie1-2. The crosses with other tolerant accessions showed that the rescue of mutant embryos defective in chloroplast translation is not limited to the Tsu-0 accession. These crosses were not examined in detail. Only a slight rescue (small globular stage) of the mutant phenotype was seen in crosses with the two sensitive accessions (Table 7, rows 3-6). Later, we found evidence of partial ACC2 function in Oy-0, and full ACC2 function in Nie1-2, which likely factors in to the slight rescue seen in these crosses.

The Suppressor Locus Maps to the *ACC2* Region of Chromosome 1

Because we believed that *ACC2* might compensate for the loss of *accD* function when chloroplast translation is blocked, we focused on *ACC2* as the possible suppressor of preglobular

Table 7. Partial Embryo Rescue in F1 Siliques from Crosses between Natural Accessions and Embryo-Defective Mutants. Adapted from Parker et al. (2014).

Mutant Allele [a]	Wild-type Accession	Siliques Screened	Seeds Screened	Percent Mutant Seeds	Percent Mutant Seeds Exhibiting Embryo Rescue	Phenotype of Rescued Embryo [b]	Average Size of Rescued Embryo (μm) [c]
emb3126-1	Tsu-0	40	1842	24.1	71.4	Most large globular; some later stages	84 ± 5.8
emb3137-1	Tsu-0	40	1939	24.3	75.4	Large globular	78 ± 4.0
emb3126-1	Oy-0	11	474	24.5	72.4	Small globular	55 ± 0.6
emb3137-1	Oy-0	20	965	26.5	75.6	Small globular	55 ± 0.4
emb3126-1	Nie1-2	11	550	24.2	Not determined [d]	Tiny globular	49 ± 1.5
emb3137-1	Nie1-2	10	491	27.1	Not determined [d]	Tiny globular	50 ± 1.1

[a] Embryo arrest in parental lines occurs at the preglobular stage.

[b] Embryo rescue was more pronounced in crosses with a spectinomycin-tolerant accession (Tsu-0) than in crosses with spectinomycin-sensitive accessions (Oy-0; Nie1-2).

[c] Mean Length ± Standard Error.

[d] Rescued mutant seeds did not differ sufficiently in size from parental mutant seeds.

Table 8. Classes of F2 Plants Identified from Tsu-0 Crosses with Mutants in a Sensitive "Nossen" Background. Adapted from Parker et al. (2014).

Parental Mutant	F_2 Class Symbol	Description of F_2 Plant Phenotype	Total Plants Identified	Total Seeds Screened	Percent Mutant Seeds	Percent Embryo Visible
emb3126-1	*SS*	No evidence of embryo rescue	21	3199	27.0	0.8
	ST	Partial rescue segregating	49	6103	26.3	76.7
	TT	Partial rescue consistent	31	7862	25.3	99.4
	WT	Wild-type plants	45	1549	0.2	-
emb3137-1	*SS*	No evidence of embryo rescue	9	1491	24.5	0.5
	ST	Partial rescue segegating	30	5259	24.7	72.8
	TT	Partial rescue consistent	19	3144	26.6	99.4
	WT	Wild-type plants	37	3993	0.9	-

arrest. Dr. Yixing Wang, a research associate in the Meinke lab, tested this hypothesis using a candidate gene approach with accession-specific PCR primers that focused initially on the *ACC2* region of chromosome 1. This approach utilized three distinct categories of F2 plants from crosses between Tsu-0 and the *emb* mutants. Yixing Wang PCR genotyped representative plants from each category for the Tsu-0 and "Nossen" alleles of *ACC2*, and showed perfect linkage between *ACC2* and the suppressor. In order to show that the Tsu-0 suppressor impacts both embryo development in the absence of chloroplast translation and seedling responses to spectinomycin, Yixing Wang PCR genotyped sensitive and tolerant F2 seedlings from the crosses between wild-type Tsu-0 and "Nossen" plants. The results showed sensitive seedlings were homozygous for the "Nossen" allele of *ACC2* while tolerant seedlings were either homozygous or heterozygous for the Tsu-0 allele, which is consistent with a dominant pattern of inheritance for the suppressor. Because these approaches are not associated with my role in this project, the details of them can be found in Parker et al. (2014).

To provide further evidence that ACC2 is the suppressor, we measured the extent of seedling growth on spectinomycin of two knockout mutants disrupted in *ACC2* in a Col-0 background. Under standard growth conditions, *acc2* mutant plants appear normal. However, mutant seedlings consistently exhibited a higher level of sensitivity to spectinomycin than wild-type (Col-0) seedlings (Figure 15). Mutant embryos homozygous for a second mutant allele of *EMB3137* (*emb3137-2*) in the Col-0 background arrest at a large globular stage of development. In order to determine if we could further impair embryo development in this mutant, we crossed *emb3137-2* with one of the *acc2* mutant lines (SALK-148966c). The results of these crosses showed 25% of mutant embryos in the F1 siliques arrested at an earlier (preglobular) stage of development (Table 9). The results from all four approaches mentioned here support the conclusion that the Tsu-0 suppressor of preglobular arrest is an allele of *ACC2*.

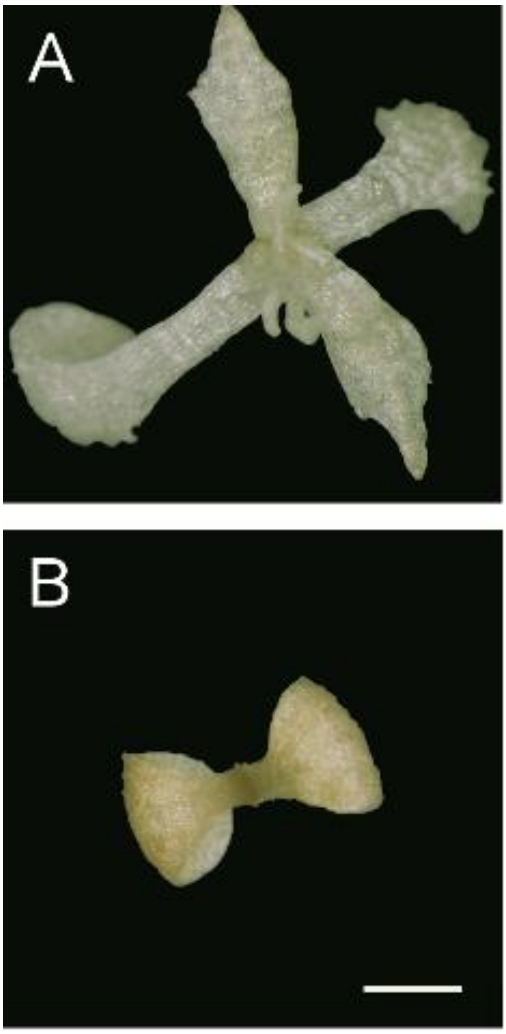


Figure 15. Spectinomycin Responses of an *acc2* Knockout Mutant Compared to the Background Accession (Col-0). A, Parental Col-0 accession. B, *acc2-1* (Salk_148966c). Bar = 1 mm. Adapted from Parker et al. (2014).

Table 9. Reduced Embryo Development in F1 Siliques from *acc2* (Col-0) Crossed with *emb3137-2* (Col-0). Adapted from Parker et al. (2014).

Cross	F_2 Seeds Screened	Percent Mutant Seeds	Mutant Seeds Screened	Percent Preglobular Embryos [a]	Average Seed Size (µm) [b]		Parental Embryo Lengths (µm) [b]	Parental Embryo Stages (%)	
					Preglobular	Parental	Average	Globular	Triangular
1	466	27.7	129	22.5	400 ± 7.2	564 ± 3.8	89 ± 2.6	90.3	9.7
2	173	32.9	57	33.3	418 ± 4.1	605 ± 4.4	92 ± 2.3	97.1	2.9
Total	639	29.1	186	25.8	407 ± 4.5	575 ± 4.2	90 ± 1.8	92.2	7.8
WT	1321	0.2	-	-	-	-	-	-	-

[a] Two classes of mutant seeds are found in F1 siliques: preglobular seeds presumed to be *acc2* homozygotes, and large seeds with globular embryos (*acc2/ACC2*; *ACC2/ACC2*) characteristic of parental *emb3137-2* lines.

[b] Mean Length ± Standard Error.

A Semidominant Enhancer Promotes Further Embryo Development in the Absence of Chloroplast Translation

Screening siliques of F2 plants from the crosses between Tsu-0 and *emb3126-1* revealed three distinct subclasses of TT plants: (1) early TT plants whose rescued mutant embryos arrested at a large globular stage of development, with very few exceeding 100 μm in diameter; (2) late TT plants whose rescued mutant embryos frequently developed beyond 100 μm, and often reached an elongated or cotyledon stage of development; and (3) intermediate TT plants whose rescued mutant embryos were a mixture of the other two classes (Table 10; Figure 16). The differences between the three enhancer classes are supported by an analysis of variance (ANOVA) on the embryo length measurements ($F = 302.9$; $p < 0.001$). These TT classes were found in a 1:2:1 ratio of early:intermediate:late plants, which is consistent with a second locus, an enhancer, that further increases the extent of embryo development in the presence of the Tsu-0 suppressor.

Curiously, TT F2 plants from crosses between Tsu-0 and *emb3137-1* could not be divided into distinct subclasses; all of these F2 plants were similar to the early TT plants of the *emb3126-1* crosses. Table 11 and Figure 17 show the differences between *emb3137-1* and *emb3126-1*. This difference is supported by a T-test on the embryo length measurments ($t = 9.8$; $p < 0.001$). The two mutant lines, *emb3137-1* and *emb3126-1*, are defective in two different chloroplast ribosomal proteins in the "Nossen" background. Because the extent of embryo development in the SALK *emb3137-2* allele was similar to other Col-0 mutants defective in chloroplast translation, we reasoned that the phenotype difference in TT plants from the *emb3126-1* and *emb3137-1* crosses was due to linkage between *EMB3137* and the enhancer locus. Yixing Wang PCR genotyped early and late TT plants from the F2 generation of the crosses between Tsu-0 and *emb3126-1* crosses for three candidate genes: *EMB3137*; *OEP80*, which is located 10 cM below *EMB3137*; and *TOC34*, which is located 10 cM above *EMB3137*. The genotyping results confirmed tight

Table 10. Enhancer Phenotype Classes of TT Plants from a Tsu-0 Cross with *emb3126-1*. Adapted from Parker et al. (2014).

Plants Analyzed [a]		Mutant Embryos Analyzed		Embryo Lengths (%)		Embryo Phenotypes (%)			
Enhancer Class	Number Screened	Number Measured	Avg. Length (µm) [b]	< 100 µm	> 200 µm	Globular	Triangular	Linear	Cotyledon
Late	26	1220	154 ± 4.0	3.6	11.3	9.5	47.8	33.9	8.8
Intermediate	46	1928	92 ± 3.9	62.7	3.5	74.6	14.9	8.7	1.8
Early	26	965	66 ± 0.9	94.8	0.1	99.4	0.5	0.1	0.0

[a] Limited to F2 plants and F3 plants derived from F2 plants assigned to the intermediate enhancer class.

[b] Mean Length ± Standard Error.

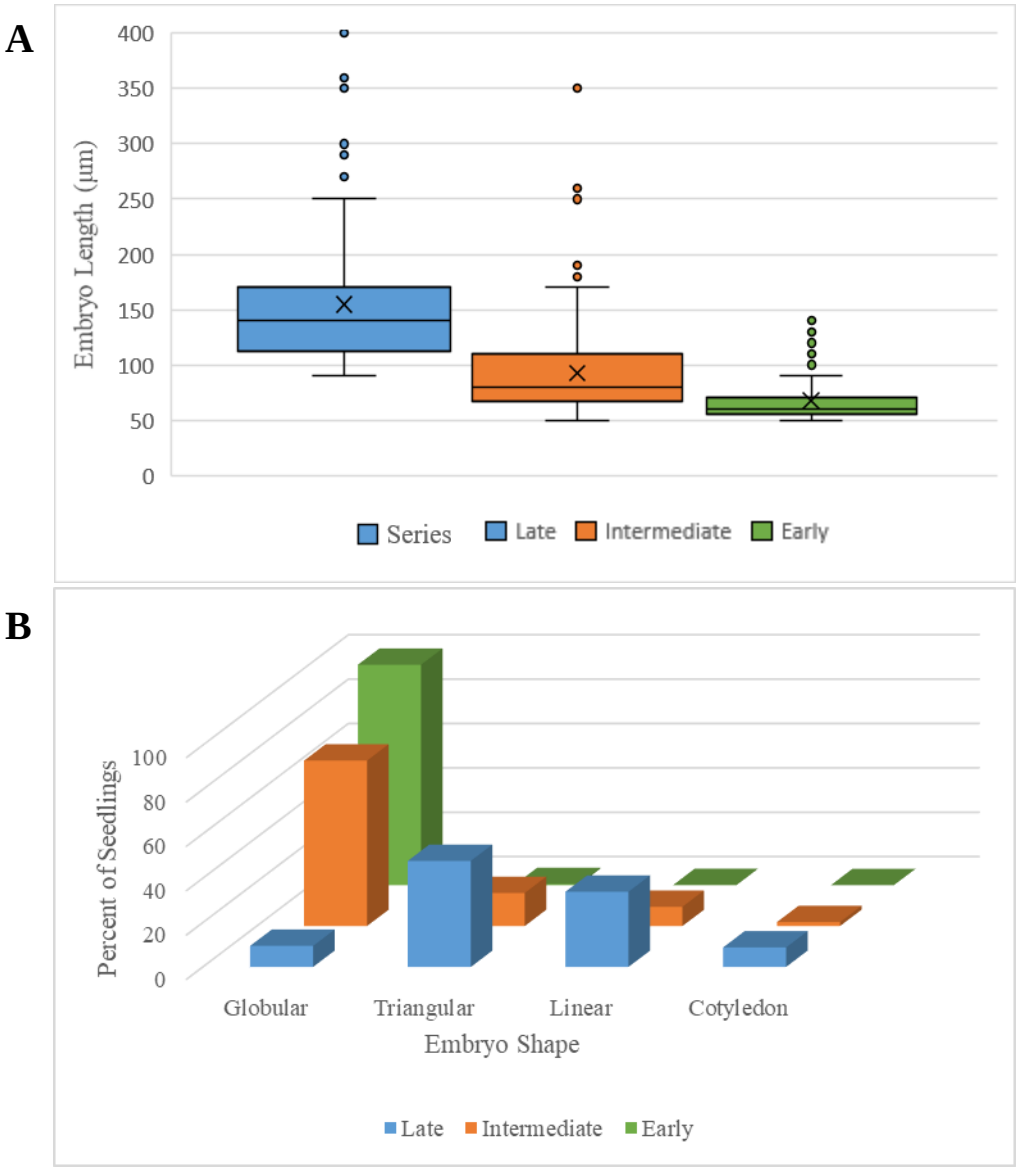


Figure 16. Enhancer Phenotype Classes of TT Plants from a Tsu-0 Cross with *emb3126-1*. A, Boxplot representing the median, 25th and 75th percentiles (interquartile range) of mutant embryo lengths. Whiskers extend to the minimum and maximum lengths (excluding outliers). Mean is denoted by the X. One extreme outlier (500 µm) for the Late enhancer class is not shown on the graph. B, Percentage of embryos in each enhancer class assigned to four phenotypic categories based on shape of the embryo: Globular, Triangluar, Linear, and Cotyledon.

Table 11. Differences in the Extent of Embryo Rescue in TT Plants from Tsu-0 Crosses with *emb3137-1* and *emb3126-1*. Adapted from Parker et al. (2014).

Mutant Allele	F_2 Plants Screened	Mutant Seeds Screened	Average Embryo Lengths (μm) [a]	Embryos Measured (%)		Embryo Phenotypes (%)			
				< 100 μm	> 200 μm	Globular	Triangular	Linear	Cotyledon
emb3137-1	20	531	77 ± 1.2	82.7	6.4	96.2	3.8	0.0	0.0
emb3126-1	31	965	99 ± 1.9	60.9	31.6	68.2	17.5	11.1	3.2

[a] Mean Length ± Standard Error.

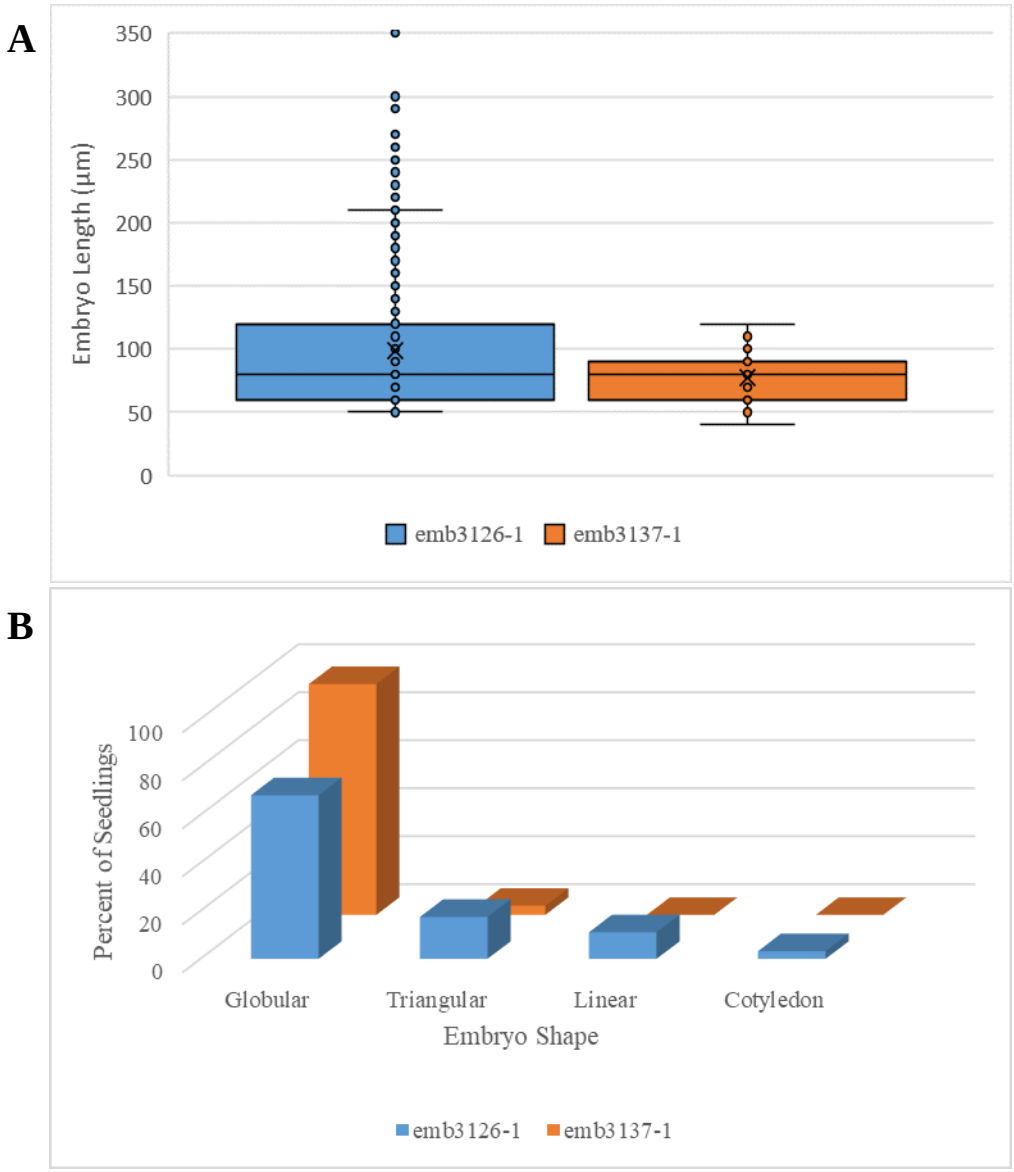


Figure 17. Differences in the Extent of Embryo Rescue in TT Plants from Tsu-0 Crosses with *emb3137-1* and *emb3126-1*. A, Boxplot representing the median, 25th and 75th percentiles (interquartile range) of mutant embryo lengths. Whiskers extend to the minimum and maximum lengths (excluding outliers). Mean is denoted by the X. B, Percentage of embryos from each cross assigned to four phenotypic categories based on shape of the embryo: Globular, Triangluar, Linear, and Cotyledon.

linkage between *EMB3137* and the enhancer, indicating that the enhancer is located near the top of chromosome 5.

We also confirmed linkage between *EMB3137* and the enhancer through crosses between Tsu-0 and two additional *emb* mutants defective in chloroplast translation: *emb1473* (Col-0), which is unlinked to *EMB3137*, and *emb3136* ("Nossen"), which is linked to *EMB3137*. Consistent with what we expected, F1 siliques of the *emb1473* crosses were similar to those seen with the *emb3126-1* crosses, with over a third of the rescued embryos developing beyond 100 μm and some embryos reaching an elongated or cotyledon stage of development (Table 12). Also as expected, F1 siliques from the crosses with *emb3136* were similar to the *emb3137-1* crosses, with rescued embryos not growing larger than 110 μm (Table 13). These results combined with Yixing Wang's PCR genotyping of the three linked loci, confirmed that the enhancer in Tsu-0 is linked to *EMB3137* and *EMB3136* near the top of chromosome 5. Further work to identify the enhancer locus has been done by Kayla Cook in our lab, and will be discussed in Chapter 6.

Additional Modifiers Increase the Frequency of Advanced Embryo Development

Analyzing the F3 siliques of progeny from late TT plants of crosses between Tsu-0 and *emb3126-1* revealed evidence of multiple modifiers that increased the frequency of advanced embryo development. These F3 plants were divided into three phenotypic categories: (1) late-advanced, where approximately 30% of the rescued embryos grew larger than 300 μm in length; (2) late-moderate, where approximately 30% of the rescued embryos grew larger than 200 μm in length but less than 300 μm; and (3) late-reduced, where approximately 85% of the rescued embryos were smaller than 100 μm when fully developed (Table 14, rows 1-3; Figure 18). The differences between the three modifier classes are supported by an analysis of variance (ANOVA) on the embryo length measurements ($F = 26.5$; $p < 0.001$). In order to determine if we could further advance embryo development from these crosses, we screened the siliques of F4

Table 12. Partial Embryo Rescue in F1 Siliques from a Tsu-0 Cross with *emb1473* (Col-0). Adapted from Parker et al. (2014).

Cross	F_2 Seeds Screened	Percent Mutant Seeds	Mutant Seeds Screened	Embryos Measured (%)		Embryo Lengths (μm)			Embryo Stages (%)			
				> 100 μm	> 150 μm	Average [a]	Min.	Max.	Globular	Triangular	Linear	Cotyledon
1	440	24.1	106	36.8	23.6	128 ± 3.6	60	470	62.2	14.2	11.3	12.3
2	711	21.7	154	37.0	21.4	117 ± 4.9	60	400	63.0	16.2	14.3	6.5
Total	1151	22.6	260	36.9	22.3	122 ± 3.0	60	470	62.7	15.4	13.1	8.8

[a] Mean Length ± Standard Error.

[b] Results are similar to *emb3126* (both genes are unlinked to the enhancer) as shown by the presence of large embryos beyond a triangular stage of development.

Table 13. Limited Embryo Rescue in F1 Siliques from a Tsu-0 Cross with *emb3136* ("Nossen"). Adapted from Parker et al. (2014).

Cross	F_2 Seeds Screened	Percent Mutant Seeds	Mutant Seeds Screened	Percent Preglobular Embryos	Average Seed Size (μm) [a]		Rescued Embryo Lengths (μm)			Rescued Embryo Stages (%)	
					Preglobular	Rescued	Average [a]	Min.	Max.	Globular	Triangular
1	988	25.8	255	20.0	370 ± 8.7	522 ± 4.2	65 ± 1.7	50	110	98.5	1.5
2	562	22.2	125	20.0	382 ± 9.7	521 ± 4.2	69 ± 1.9	50	110	99.0	1.0
Total	1550	24.5	380	20.0	374 ± 6.6	521 ± 3.0	66 ± 1.3	50	110	98.7	1.3

[a] Mean Length ± Standard Error.

[b] Results are similar to *emb3137* (both genes are linked to the enhancer) as shown by the absence of large embryos beyond a triangular stage of development.

Table 14. Modifier Phenotype Classes of Late TT Plants from a Tsu-0 Cross with *emb3216-1*. Adapted from Parker et al. (2014).

Plants Analyzed			Mutant Embryos Analyzed		Embryo Lengths (%)			Embryo Phenotypes (%)			
Modifier Class	Plant Generation	Number Screened	Number Measured	Avg. Length (µm) [a]	< 100 µm	> 200 µm	> 300 µm	Globular	Triangular	Linear	Cotyledon
Late; Advanced[a]	F_3	4	245	256 ± 12.1	0.0	61.5	30.1	0.0	10.5	45.8	43.9
Late; Moderate[a]	F_3	12	474	184 ± 11.1	0.6	25.9	6.1	5.4	20.2	55.2	19.2
Late; Reduced[a]	F_3	8	435	146 ± 6.7	3.4	6.6	1.0	14.7	38.8	41.0	5.5
Late; Advanced; Late[b]	F_4	3	134	345 ± 10.7	0.0	94.5	60.9	0.0	0.0	29.3	70.7
Late; Advanced; Moderate[b]	F_4	11	569	254 ± 12.6	0.0	64.7	28.3	0.0	4.3	52.9	42.8

[a] Mean Length ± Standard Error.

[b] Progeny plants from the "late" class of F2 plants homozygous Tsu-0 for the suppressor and enhancer.

[c] Progeny plants from the "late; advanced" class of F3 plants homozygous Tsu-0 for the suppressor and enhancer.

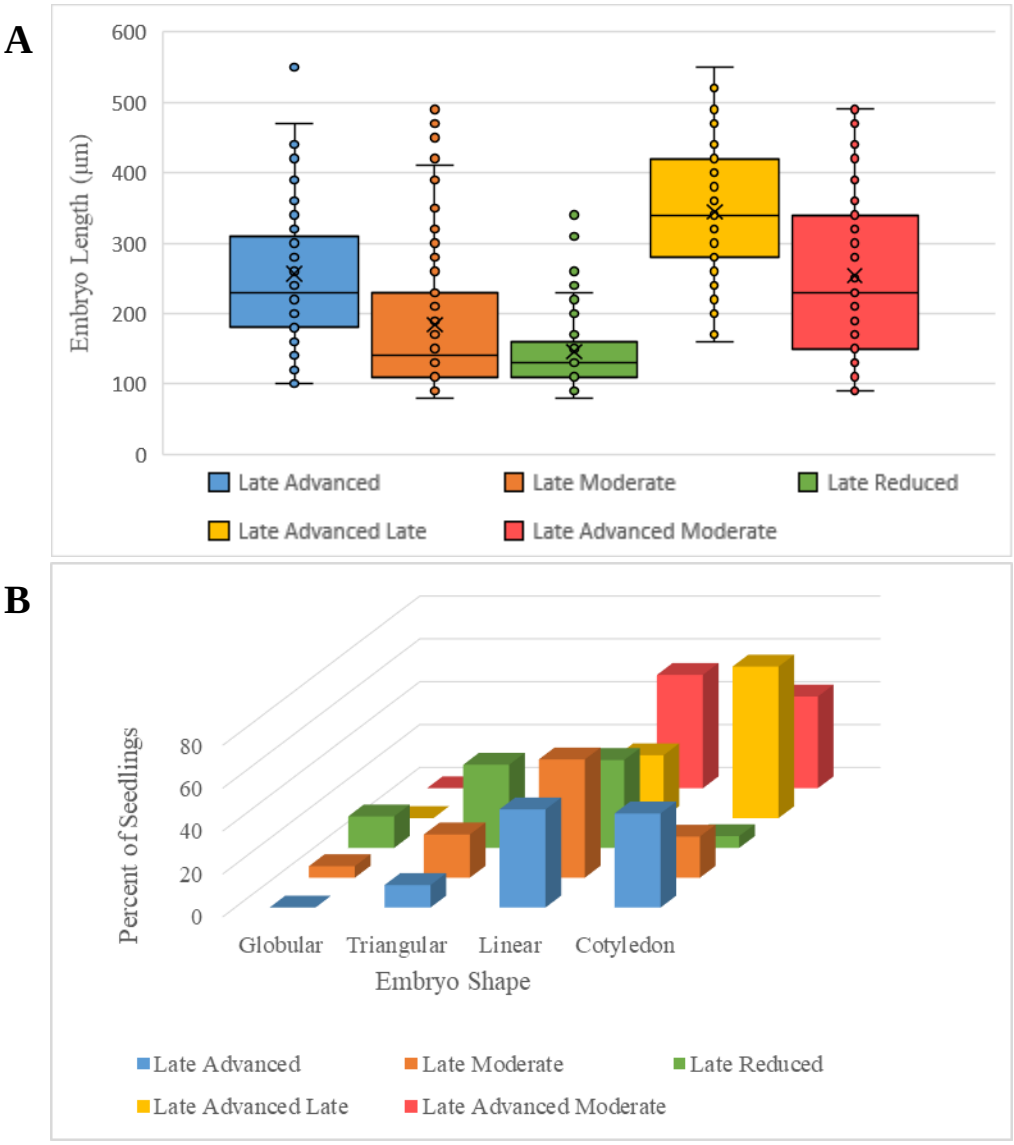


Figure 18. Modifier Phenotype Classes of Late TT Plants from a Tsu-0 Cross with *emb3126-1*. A, Boxplot representing the median, 25th and 75th percentiles (interquartile range) of mutant embryo lengths. Whiskers extend to the minimum and maximum lengths (excluding outliers). Mean is denoted by the X. B, Percentage of embryos from each modifier class assigned to four phenotypic categories based on shape of the embryo: Globular, Triangluar, Linear, and Cotyledon.

progeny from late-advanced plants. A few plants, classified as late-advanced-late, contained more than 60% of the rescued embryos larger than 300 μm, and all of the rescued embryos reached the elongated or cotyledon stages of development (Table 14, row 4). The remaining F4 plants were classified as late-advanced-moderate, and resembled the late-advanced F3 plants (Table 14, row 5). T-tests showed that the difference between the late-advanced-late and late-advanced-moderate modifier classes is statistically significant ($t = 5.5$; $p < 0.001$), whereas there is no significant difference between the late-advanced-moderate F4 plants and the late-advanced F3 plants ($t = -0.1$; $p = 0.4$). Analysis of siliques from F5 progeny from the late-advanced-late plants revealed no detectable difference in the extent of embryo rescue from the F4 generation. No fully developed, albino embryos were found among the advanced embryos screened. Details of the entire collection of plants screened for mapping the enhancer locus and identifying additional modifiers are presented in Appendices C and D respectively.

DISCUSSION

In order to identify the genes impacting phenotype differences between accessions when chloroplast translation is blocked, we utilized crosses between the tolerant Tsu-0 accession and *emb* mutants defective in chloroplast translation in the sensitive "Nossen" accession. Screening the extent of embryo rescue in these crosses gave us a more accurate system for gene identification than the seedling crosses discussed in Chapter 3, where it was difficult to classify borderline seedlings. With this approach, we identified *ACC2* as a single, dominant suppressor of the preglobular phenotype of the RIKEN *emb* mutants. This suppressor is able to rescue embryo development to a late globular stage. We also found evidence of an unlinked enhancer of the suppressor that allows embryos to develop beyond the globular stage, and additional modifiers that increase the frequency of embryos at the most advanced stages of development. These additional modifiers can also advance slightly the development of embryos when the enhancer is

not present. The effects of the Tsu-0 suppressor, enhancer, and modifier alleles are summarized in Figure 19, and examples of arrested embryo phenotypes are shown in Figure 20.

Even in the most advanced progeny examined from the crosses between Tsu-0 and *emb3126-1*, we never found a fully rescued, albino embryo. This was not surprising given that tolerant accessions grown on spectinomycin were not as fully developed as most albino mutants defective in photosynthesis alone. This is further evidence that *accD* is not the only gene in the chloroplast genome required for proper plant development. As discussed in Chapter 3, *ycf1*, *ycf2,* and *clpP1* potentially play important roles in later stages of seedling and embryo development in Arabidopsis.

Through PCR genotyping and analysis of crosses, we have identified the Tsu-0 suppressor as *ACC2* and have mapped the enhancer close to the top of chromosome 5 (linked to *EMB3137*). However, we have not identified specific genes that encode the enhancer and additional modifier proteins. One potential role for the enhancer is as a critical component of the TIC/TOC chloroplast protein import system, specifically involved in the import of ACC2 into the stroma of the chloroplast. However, disruption of this protein must not affect the import of other chloroplast-localized proteins. In this scenario, the additional Tsu-0 modifiers could encode other components of the TIC/TOC protein import system. Candidate genes for the additional modifiers include Toc132/Toc120, which are thought to be involved in recognizing and guiding housekeeping proteins through the outer membrane (Kubis et al., 2004; Inoue et al., 2010); and Tic20-IV, which is believed to be the main channel protein for some of the housekeeping proteins through the inner membrane (Hirabayashi et al., 2011). However, there are no promising candidate loci with such functions in the enhancer region on chromosome 5.

Because ACC2 is a large protein that must be imported into the chloroplast, a second possible role for the Tsu-0 enhancer is as a chaperone protein involved in the folding, guiding, or

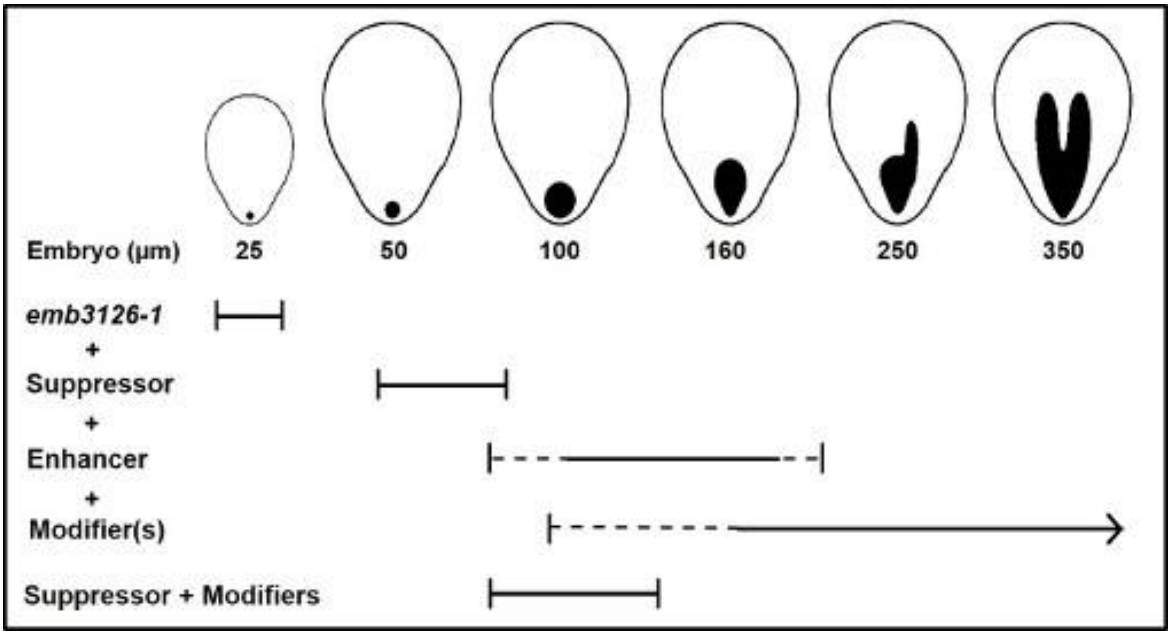


Figure 19. Combined Effects of the Tsu-0 Suppressor, Enhancer, and Modifier(s) on Seed and Embryo Rescue in *emb3126-1*. Ellipses represent mutant seeds, filled images depict mutant embryos, and bars define the stage of arrest. Adapted from Parker et al. (2014).

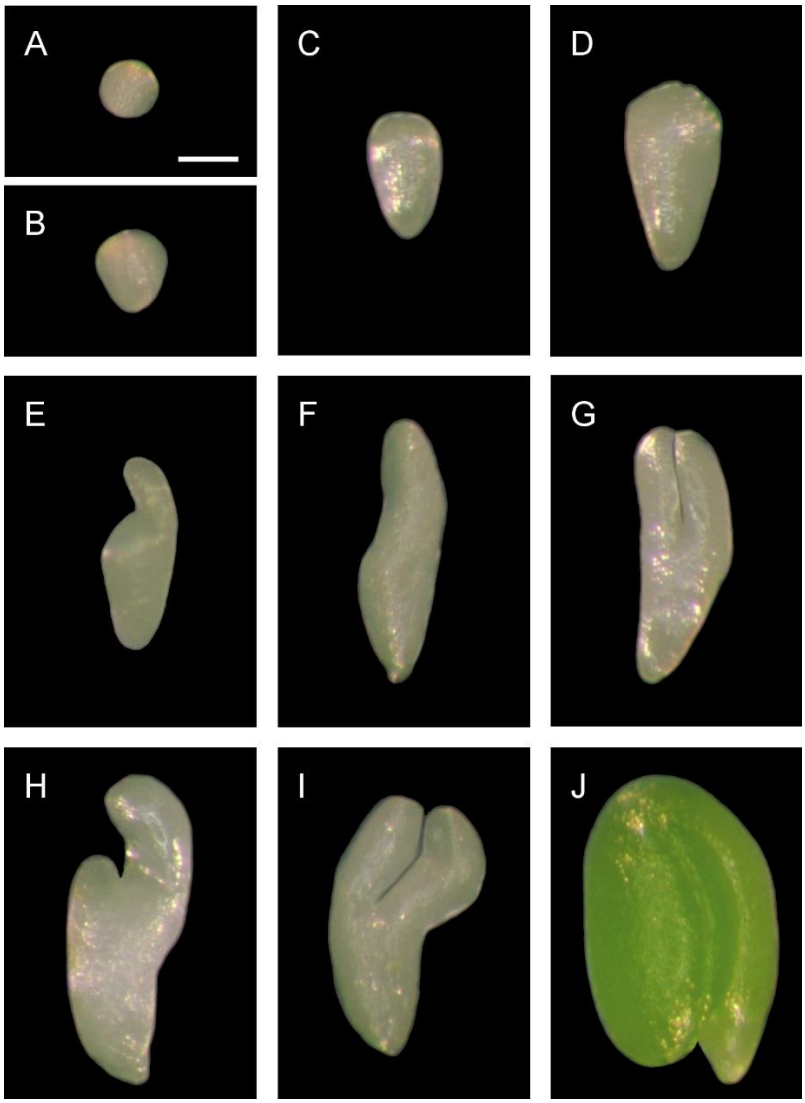


Figure 20. Examples of Embryos in Siliques of Plants Homozygous for the Tsu-0 Suppressor. A, Late globular embryo. B, Triangular embryo. C and D, Elongated linear embryos. E to I, Cotyledon stage embryos with one or two cotyledons present. J, Sibling wild-type embryo. Bar = 100 µm. Adapted from Parker et al. (2014).

stabilization of ACC2. In this scenario, the additional Tsu-0 modifier proteins could either be components of the TIC/TOC protein import system, or additional chaperone proteins. Candidate genes for the additional modifiers include Hsp70 and members of the 14-3-3 protein family, which are thought to work together in the cytosol to guide precursor proteins to the chloroplast (May and Soll, 2000; Flores-Pérez and Jarvis, 2013); Hsp93, cpHsp70 and Hsp90C, which are thought to function in the stroma of the chloroplast to stabilize and guide proteins through the inner membrane (Kovacheva et al., 2007; Inoue et al., 2013; Shi and Theg, 2013); and Tic22, which is believed to guide precursor proteins across the intermembrane space (IMS) between the TOC and TIC import complexes (Kouranov et al., 1998; Shi and Theg, 2013). However, once again: no promising chaperone genes can be found in the enhancer region.

We have shown here that the additional Tsu-0 modifiers can function to advance somewhat the development of embryos independent from the enhancer, which means these additional modifiers could have a separate function. Some modifiers could potentially function in partial compensation for the loss of *ycf1*, *ycf2*, and *clpP1* in early stages of embryo development. A candidate gene approach to look at potential modifiers is described in Chapter 6.

After determining that *ACC2* impacts the phenotypic differences between Arabidopsis accessions when chloroplast translation is blocked, we decided to look at how changes in *ACC2* increased the tolerance of some accessions to spectinomycin. We first thought that *ACC2* might be overexpressed in tolerant accessions, which would increase the amount of *ACC2* transcript and possibly the amount of ACC2 protein present in the chloroplast. However, RT-qPCR experiments by Yixing Wang showed no significant difference in the amount of *ACC2* transcript found in tolerant and sensitive accessions (Parker et al., 2014). We then focused on the protein sequence of ACC2 thinking that a change in the transit peptide could increase the localization or amount of protein taken into the chloroplast, or a mutation in the protein sequence could increase the activity of ACC2 or increase the interactions of ACC2 with chaperone proteins. Around the time we

began to look at the sequenced genomes available through the 1001 Genomes Project (The 1001 Genomes Consortium, 2016), Yixing Wang sequenced the *ACC2* gene from the sensitive "Nossen" accession and found a nonsense mutation in the middle of the gene. We then changed our approach from looking at what causes tolerance to a loss of chloroplast translation to looking at what can cause sensitivity. Chapter 5 in this dissertation discusses the diversity of *ACC1* and *ACC2* mutations found in natural Arabidopsis accessions.

CHAPTER V

A VARIETY OF *ACC2* MUTATIONS ARE FOUND IN NATURAL ACCESSIONS OF ARABIDOPSIS

INTRODUCTION

Sequencing the *ACC2* gene from the "Nossen" accession by Yixing Wang changed our perspective on the phenotypic differences found between accessions when chloroplast translation is blocked. Previously, we looked for mutations in *ACC2* that increased the tolerance of an accession to spectinomycin. After analyzing the *ACC2* sequence from the sensitive "Nossen" accession, we began to look for other changes in *ACC2* that caused sensitivity to a loss of chloroplast translation. We combined our experimental system to evaluate the level of ACC2 function using sensitivity to spectinomycin with the genome sequence data from the 1001 Genomes Project (http://signal.salk.edu/atg1001; The 1001 Genomes Consortium, 2016) to analyze the relationship between genotype and phenotype within the ACCase class of proteins, which are essential for eukaryotic fatty acid biosynthesis. The *ACC2* experimental system in Arabidopsis provides a unique opportunity to look at the deleterious effects of different types of mutations on an essential class of proteins with implications for agriculture and human health.

This chapter describes various *ACC2* mutations found in sensitive accessions of Arabidopsis. Utilizing the 1001 Genomes Project sequences, we used two methods to look at what determines sensitivity in Arabidopsis accessions. We first used the forward genetic approach described in Chapter 3 to expand our list of sensitive accessions by testing 100 random accessions from the 1001 Genomes Project. The second method, a reverse genetic approach, focused on known variation in *ACC2* sequence among 855 sequenced accessions. Rather than testing all of the variants found, we utilized sequence conservation from an alignment of 667 eukaryotic ACCases to identify conserved regions where variation in the protein sequence would most likely lead to sensitivity. We also tested accessions with variants in the transit peptide at the N-terminus of ACC2. Among the sensitive accessions discovered through both of these approaches, we found that sensitivity could be caused by nonsense mutations, frameshifts, defects in RNA splicing recognition sites, large deletions or sequence rearrangements, small deletions, and missense mutations in residues that are likely essential for ACC2 function. Confirmation that the mutations found in *ACC2* cause sensitivity through reduced or eliminated protein function was done using two approaches: (1) crossing sensitive accessions with the tolerant Tsu-0 accession in order to show linkage between the sensitive phenotype and the *ACC2* genotype; and (2) genetic complementation tests between each sensitive accession of interest and informative *acc2* and *tic20-iv* knockout mutants.

Through the analysis of all natural variation of ACC1 and ACC2 protein sequences among 855 sequenced accessions, we identified 339 variant residues (15% of all residues in ACCase). Of these variants, five significantly reduce or eliminate protein function, 18 partially reduce function, and 316 have no significant effect on ACCase function. Most of the data presented in this chapter have been published (Parker et al., 2014; 2016). Exceptions include results from the Qar-8a, Ts-1, and Etna-2 crosses with the knockout mutants, which were obtained after publication.

MATERIALS AND METHODS

ACC2 Variation in Arabidopsis Accessions

ACC2 protein sequences from 855 natural accessions of Arabidopsis were obtained from the Salk Institute 1001 Genomes Project website (http://signal.salk.edu/atg1001; Appendix E). These sequences were entered into an Excel spreadsheet to track variation in the amino acid residues. The spreadsheet was organized so that each row consisted of the full ACC2 protein sequence from one accession while each column displayed the amino acid at a specific residue in the sequence. In addition, the ACC2 sequences of "Nossen" and Sav-0 were added to the spreadsheet for a total of 857 sequences. A list of the formulas used with this spreadsheet is found below. Variation was tracked using formula A, which counts the number of lines (accessions) that contain the same amino acid as the Col-0 sequence, which is used as a template. For residues where variation is found, formula B was used to identify the most common amino acid at that residue, and formula C to count the number of accessions with that amino acid. Similarly, formula D was used to identify the least common amino acid at a residue, and formula E to count the number of accessions with that amino acid. If additional amino acid variation was present at a residue, then the different amino acids were identified visually, and formula F was used to count the number of accessions with that amino acid.

A. =COUNTIF(B2:B858,CONCATENATE("=",B862))
B. =INDEX(E2:E858,MATCH(MAX(COUNTIF(E2:E858,E2:E858)),COUNTIF(E2:E858,E2:E858),0))
C. =COUNTIF(E2:E858,E866)
D. =INDEX(E2:E858,MATCH(MIN(COUNTIF(E2:E858,E2:E858)),COUNTIF(E2:E858,E2:E858),0))
E. =COUNTIF(E2:E858,E868)
F. =COUNTIF(E2:E858,E870)

Brassicaceae ACCase Sequence Analyses

For the comparison of ACC1 and ACC2 sequences, determination of K_a (nonsynonymous nucleotide substitutions) to K_s (synonymous substitutions) ratios, which is used to analyze the selection pressure on a gene, and analysis of *ACC2* Intron 6, genomic sequences for six members of the Brassicaceae were downloaded from the Phytozome (www.phytozome.net; Goodstein et al., 2012) and CoGe (www.genomevolution.org/CoGe/; Lyons et al., 2008) websites: Arabidopsis, *Arabidopsis lyrata* (Hu et al., 2011), *Brassica rapa* (Cheng et al., 2011), *Capsella rubella* (Slotte et al., 2013), *Leavenworthia alabamica*, and *Sisymbrium irio* along with *Theobroma cacao* (Motamayor et al., 2013). Appendix F lists details of the sequences used for these comparisons. K_a/K_s ratios were calculated with the coding sequences using MEGA version 6 (Tamura et al., 2013). These genomic sequences were aligned using ClustalW2 (Larkin et al., 2007).

Eukaryotic, Homomeric ACCase Sequence Alignments

In order to identify conserved amino acid residues potentially important for function of ACCase proteins, we created three alignments of protein sequences. The first utilized 20 ACC1 and ACC2 sequences from model organisms: Arabidopsis (2), *B. rapa* (2), *Medicago truncatula* (1), *Triticum aestivum* (2), *Zea mays* (2), *Homo sapiens* (2), *Mus musculus* (2), *Danio rerio* (2), *Drosophila melanogaster* (1), *Saccharomyces cerevisiae* (2), *Schizosaccharomyces pombe* (1), and *Neurospora crassa* (1). Appendix G lists details of the sequences used for this alignment. These protein sequences were aligned using ClustalW2 (Larkin et al., 2007). In several of these protein sequences, we found small gaps that we believed to be annotation errors when translating the genomic data. In order to fill in the amino acids from these gaps, we utilized the original genomic data.

In order to increase the number of eukaryotic homomeric ACCase sequences in the multi-

kingdom alignment, 744 protein sequences were selected from the Pfam database based on the presence of a large central domain, which is unique to eukaryotic homomeric ACCases (http://pfam.xfam.org/family/PF08326). From this group, four sequences from *Caenorhabditis elegans* lacked the lysine residue that is required for biotin to bind to the protein and were subsequently removed from the list. Five more sequences were also removed because they were fragmented, and 104 bacterial sequences were removed because the large central domain is unique to eukaryotes. In order to increase the number of plant sequences present, 36 plant sequences were identified through BLAST searches using both ACC1 and ACC2 Arabidopsis sequences. The final list of sequences in this expanded multi-kingdom alignment totaled 667: 198 animal, 139 plant, 276 fungal, and 54 others such as algae and protozoa (Appendix H). The actual percentage of conservation for some residues may be slightly higher than calculated due to the presence of small gaps in some sequences that are likely annotation errors. All 667 sequences were aligned using the MUSCLE program (Edgar, 2004) through Jalview 2.8.2 analysis workbench (Waterhouse et al., 2009). The 139 plant sequences were also aligned separately using the same MUSCLE program.

RESULTS

Null Mutations in *ACC2* Are Found Among Natural Accessions

Following the discovery by Yixing Wang of the nonsense mutation in *ACC2* from the sensitive "Nossen" accession, we wondered whether the mutation in "Nossen" was unique or if there were other natural accessions that contained null mutations in *ACC2*. To determine this, we examined *ACC2* sequences from 855 accessions available through the 1001 Genomes Project for additional examples of nonsense mutations and other types of null mutations. Table 15 lists the accessions we identified with various null mutations in *ACC2*. All of the variants listed were

Table 15. *ACC2* Null Mutations Identified in Sequenced Accessions of Arabidopsis. Adapted from Parker et al. (2016).

Mutation Class	Accession	Reported Country of Origin	Mutation [a]	Mutation Location	Spectinomycin Response		
					Category	Score [b]	Seedlings
Nonsense	Kb-0	Germany	Y753X	Exon 17	Sensitive	1.4	73
	Kl-5	Germany	Y753X	Exon 17	Hypersensitive	1.1	76
	"Nossen"	Uncertain	R865X	Exon 19	Sensitive	2.3	571
	Blh-1	Czech Republic	K1225X	Exon 26	Sensitive	1.3	71
Frameshift	Ip-Alo-0	Portugal	1171fs	Exon 25	Hypersensitive	1.1	51
	Ip-Vin-0	Spain	1171fs	Exon 25	Hypersensitive	1.2	33
	Lu3-30	Germany	2020fs	Exon 31	Hypersensitive	1.1	54
	Lu4-2	Germany	2020fs	Exon 31	Hypersensitive	1.3	55
Splicing	Gn-1	Germany	GT…TG	Intron 10	Hypersensitive	1.1	83
	"Gn2-3"	Germany	GT…TG	Intron 10	Hypersensitive	1.1	191
	Wl-0	Germany	GT…GG	Intron 19	Sensitive	1.4	79
	Spro-2	Sweden	TT…AG	Intron 29	Sensitive	1.3	80
	Ste-2	Sweden	TT…AG	Intron 29	Hypersensitive	1.1	83
	Ste-3	Sweden	TT…AG	Intron 29	Hypersensitive	1.0	82
	Vimmerby	Sweden	TT…AG	Intron 29	Hypersensitive	1.0	67
Rearrangement	Ob-0	Germany	Unresolved	Exon 32	Hypersensitive	1.2	74
	Old-0	Germany	Unresolved	Exon 32	Hypersensitive	1.2	75
Small Deletion	Ip-Ber-0	Spain	Deletion (23 bp)	Intron 17; Exon 18	Sensitive	1.3	49

[a] All variants except details of Ob-0 and Old-0 rearrangements were confirmed by Sanger sequencing.

[b] Higher scores reflect increasing levels of tolerance; these scores were among the lowest of all accessions evaluated.

confirmed by Yixing Wang using Sanger sequencing except for the large deletions or chromosomal rearrangements found in Ob-0 and Old-1. I screened seedlings from each of these accessions on spectinomycin, and found most of them to be highly sensitive. Including the mutation in "Nossen", we identified four different nonsense mutations: Y753X (Kb-0 and Kl-5), R865X ("Nossen"), K1225X (Blh1-1), and Q2325X (Hod). Three of these mutations result in truncated ACC2 proteins that are missing over 1,000 amino acids from the C-terminus. The fourth mutation (Q2325X) results in a truncated protein that is missing only 30 amino acids from the C-terminus. The seedling phenotypes of these five accessions on spectinomycin are consistent with the severity of the protein truncation. Seedlings from Kb-0, Kl-5, "Nossen", and Blh1-1 are sensitive to a loss of chloroplast translation while seedlings from Hod are phenotypically intermediate.

We identified two different frameshift mutations caused by single nucleotide deletions. One of these (1171fs1190X) was found in the central domain of the IP-Alo-0 and IP-Vin-0 accessions, and resulted in a downstream nonsense mutation and removal of 1,165 amino acids from the C-terminus. The other frameshift mutation (2020fs2021X) was found in the carboxyltransferase α-subunit of the Lu3-30 and Lu4-2 accessions, and resulted in an immediate nonsense mutation and a truncated ACC2 protein missing 334 C-terminal amino acids. By comparing RNA splicing recognition sites for the introns in *ACC2*, we identified two mutations in splice acceptor sites (intron 10 and intron 19) and one in a splice donor site (intron 29). Both of the altered splice acceptor sites result in a 10-nucleotide deletion and a frameshift mutation. In the Gn-1 accession, there is an AG → TG substitution in the acceptor site of intron 10. The Wl-0 accession contains an AG → GG substitution in the acceptor site of intron 19. Four accessions (Spro-2, Ste-2, Ste-3, and Vimmerby) contain the same GT → TT substitution in the donor site of intron 29. Using Spro-2 to represent the group, Yixing Wang showed that this mutation results in a mixture of defective *ACC2* transcripts that include all or some of intron 29.

Large deletions, or possibly chromosomal rearrangements, were identified at the C-terminal end of the *ACC2* sequences for Ob-0 and Old-1. Yixing Wang confirmed that both accessions are missing exon 31, but she was unable to resolve the exact nature of the defect. A small deletion of 23-nucleotides was identified in IP-Ber-0. This deletion removes the end of intron 17 and the beginning of exon 18. Using RT-PCR, Yixing Wang showed that this deletion results in a mixture of *ACC2* transcripts that encode a variety of defective and truncated proteins.

The Structure of *ACC2* Sequences Varies Within the Brassicaceae

In order to get a broader perspective on natural variation in the structure and function of *ACC2*, we compared *ACC1* and *ACC2* sequences from members of the Brassicaceae whose genomes have been sequenced (Table 16; Figure 21). Similar to Arabidopsis, tandem gene duplications of *ACC1* and *ACC2* are present in *Arabidopsis lyrata*, *Capsella rubella*, and *Eutrema parvulum*, while unlinked copies are found in *Brassica rapa*. *ACC1* and *ACC2* are also found in the genomes of *Sisymbrium irio* and *Leavenworthia alabamica*, but it is unclear whether they are linked or unlinked. In the *ACC2* sequence of *L. alabamica*, a nonsense mutation is found in the third exon. We found no evidence of *ACC2* in the sequenced genomes of *Aethionema arabicum* and *Boechera stricta*. Sequence comparison between *ACC1* and *ACC2* of six Brassicaceae members revealed more sequence variation in *ACC2* than in *ACC1*. Only two amino acid residues differed in ACC1 among the six Brassicaceae members while 17 residues differed in ACC2. Comparing the frequencies of synonymous (K_s) to nonsynonymous (K_a) substitutions using these sequences showed a slight relaxation of purifying selection in *ACC2* when compared to *ACC1* (K_a/K_s ratios: *ACC1*, 0.08; *ACC2*, 0.20).

In addition to increased variation in *ACC2* sequence when compared to *ACC1*, the *ACC2* gene in Arabidopsis contains a large intron (2.5 kb) that interrupts the biotin carboxylase domain, which could diminish production of ACC2 by reducing the levels of the full-length *ACC2* mRNA

Table 16. Variation in Brassicaceae *ACC1* and *ACC2* Sequences.

Distribution of *ACC1* and *ACC2*	Example Species
Tandem Duplication	*Arabidopsis thaliana* *Arabidopsis lyrata* *Capsella rubella* *Eutrema parvulum*
Non-Tandem Duplication	*Brassica rapa*
ACC2 Present; Linkage Unknown	*Sisymbrium irio* *Leavenworthia alabamica*
ACC2 Nonsense Mutation	*Leavenworthia alabamica*
ACC2 Absent	*Aethionema arabicum* *Boechera stricta*

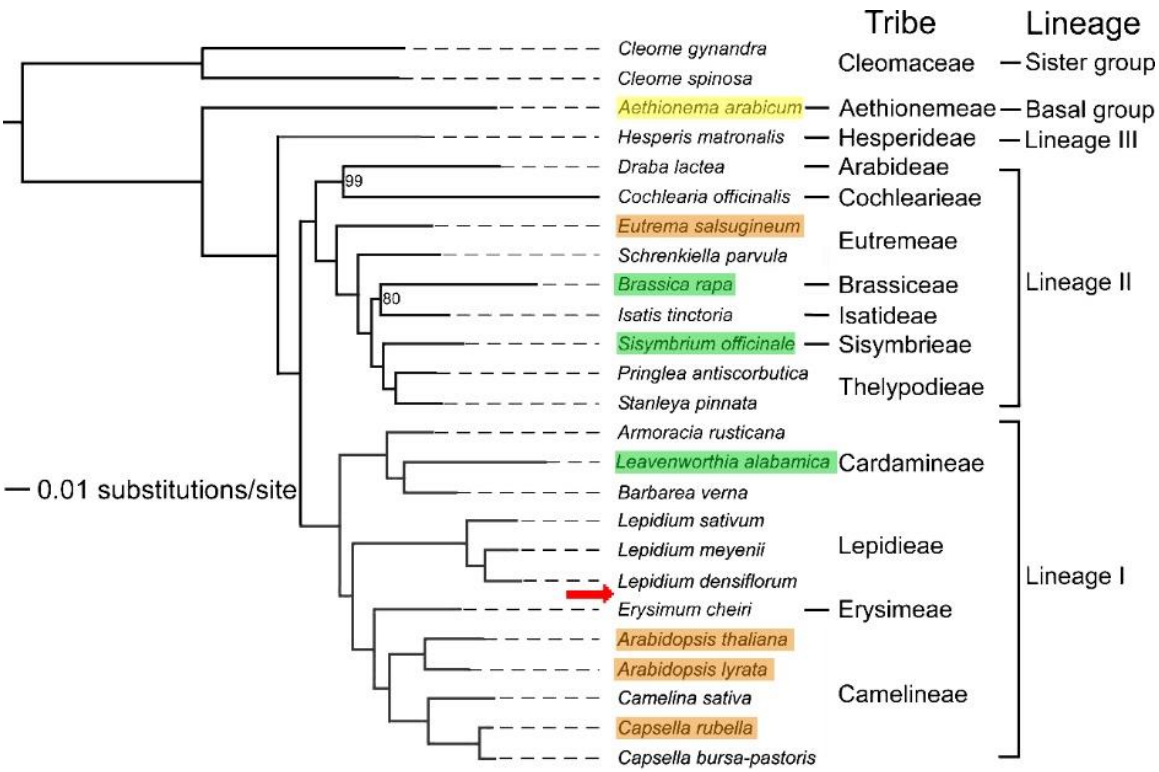


Figure 21. Brassicaceae Phylogeny. Yellow, *ACC2* is missing from the nuclear genome; Orange, *ACC2* is present as a tandem duplication; Green, *ACC2* is present as a non-tandem duplication or the location of the duplication is unknown; Red Arrow, location of *Boechera stricta* in the phylogeny, which is missing *ACC2*. Adapted from Kagale et al. (2014).

transcribed. Variation in the length of this intron (#6) can be seen among the 855 sequenced accessions. Within other members of the Brassicaceae, this intron is not as large as in Arabidopsis. A sizeable intron (810 to 1573 bp) can be found in *ACC2* sequences of *A. lyrata*, *C. rubella*, *S. irio*, and *Camelina sativa*. Intron 6 in *B. rapa* and *L. alabamica* is similar in length (72 to 217 bp) to the intron found in *ACC1* of Arabidopsis. Another gene in Arabidopsis, At3g52700, which encodes a protein of unknown function, contains an intron that matches 1 kb from the middle of the *ACC2* intron. The intron in *ACC2* seems to contain a degenerate helitron transposon that is nested within a MULE (Mutator-Like) element (Thomas Bureau, personal communication). The matching region in At3g52700 appears to be a related helitron transposon. These results provide evidence of multiple gene insertions that targeted *ACC2* following the initial duplication of *ACC1*.

According to the locus page for *ACC2* at TAIR (http://www.arabidopsis.org/), a second gene model predicts a small transcript that terminates at Intron 6, which would encode a truncated ACC2 protein missing part of the biotin carboxylase domain along with all other domains. Yixing Wang confirmed the presence of this shorter transcript in Col-0. Believing that the large size of Intron 6 might decrease the amount of full-length *ACC2* transcript produced, which could affect the tolerance of an accession to a loss of chloroplast translation, Yixing Wang compared levels of the short transcript between tolerant and sensitive accessions. No evidence was found that increased levels of the short transcript affected the tolerance or sensitivity of an accession.

Conservation Found in Alignments of Eukaryotic, Homomeric ACCase Sequences

After identifying a number of null mutations that eliminate ACC2 protein function, we began to look for conserved regions in the *ACC2* sequence that when altered might affect protein function. We first approached this using an alignment of 20 homomeric, eukaryotic ACC1 and

ACC2 protein sequences from model organisms, including nine plant sequences. Using this alignment, we identified 416 amino acid residues (out of 2,355 total) that were perfectly conserved across all 20 sequences. In order to narrow down this list of conserved residues to those most likely to be essential for ACC2 function, we expanded our multi-kingdom alignment to include 667 homomeric, eukaryotic ACCase protein sequences, including 139 plant sequences. Using this new alignment, we identified 526 amino acid residues that are more than 90% conserved across all sequences, and 222 residues that are at least 99% conserved. In addition to the 667-sequence multi-kingdom alignment, we aligned 139 homomeric, plant ACCase sequences. In this plant alignment, we identified 1196 amino acid residues that are more than 90% conserved and 698 residues that are at least 99% conserved. These alignments, especially the 667-sequence multi-kingdom alignment, were used in both the forward and reverse genetic approaches to identify ACC2 amino acid residues that are likely essential for protein function. The percent conservation for all amino acid residues in ACC2 is shown in Appendix I, along with variation in ACC1 and ACC2 protein sequences for the 855 Arabidopsis accessions.

Figure 22 and Appendix J show the mutational landscape of homomeric ACCases in Arabidopsis. Represented in these images are locations for mutations in Arabidopsis *ACC1* and *ACC2* sequences that have been either induced or found in natural accessions, highly conserved residues (>95%) from the 667-sequence multi-kingdom alignment of ACCases, and all of the natural variation found in both *ACC1* and *ACC2* of sequenced Arabidopsis accessions. Appendix K provides details on informative variants and residues in *ACC2*, including the mutations shown in the mutational landscape, and mutations in other model organisms.

Sensitive Accessions Highlight Conserved Residues Likely to be Essential

Utilizing data from the forward genetic screens of Arabidopsis accessions on spectinomycin described in Chapter 3, we identified three hypersensitive and 22 sensitive

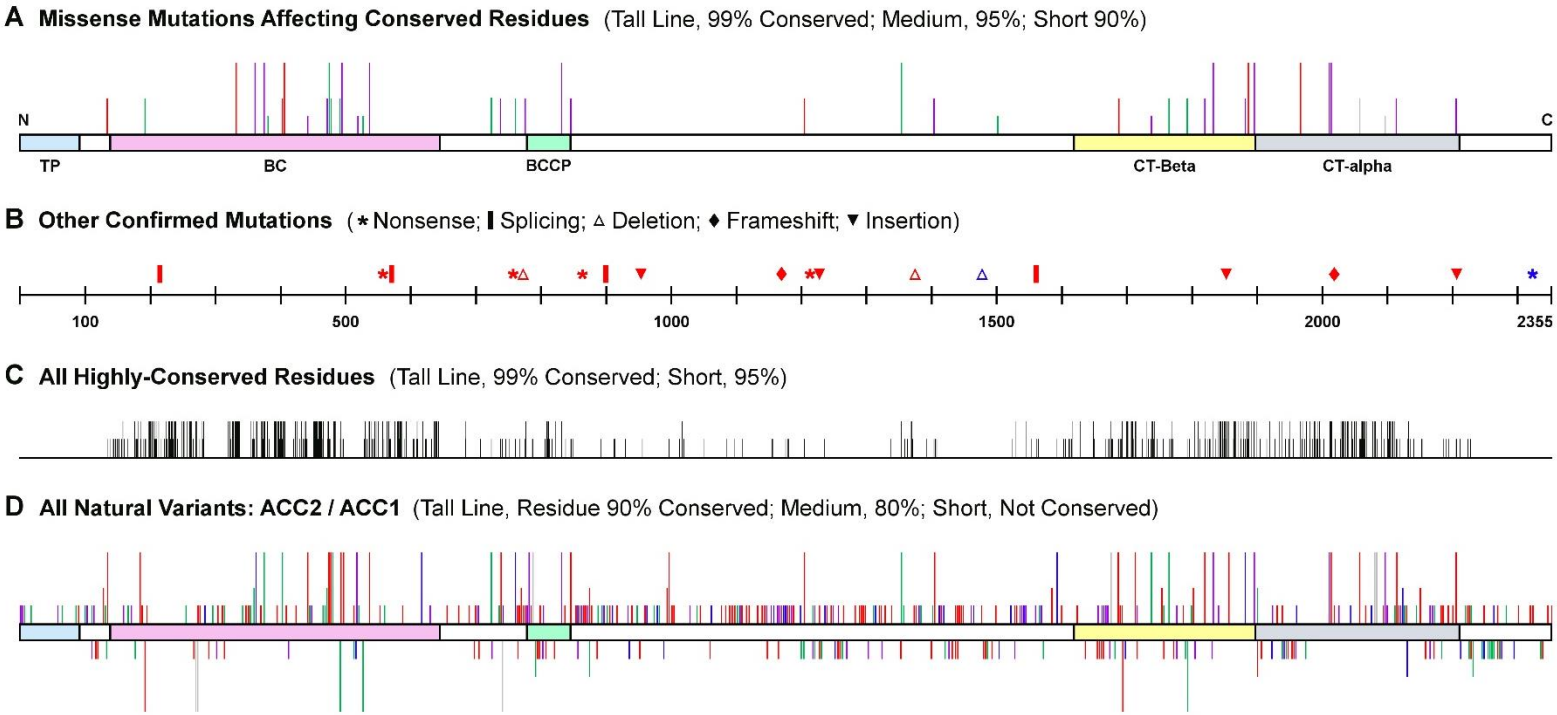


Figure 22. Mutational Landscape of Homomeric ACCase in Arabidopsis. Conservation percentages are based on the multi-kingdom alignment of 667 ACCase protein sequences. A, Induced and natural variants in ACC1 and ACC2. Red bars, deleterious or likely deleterious variants; purple bars, possibly deleterious; green bars, not deleterious or likely not deleterious; gray bars, variants of unknown significance. TP, transit peptide domain; BC, biotin carboxylase; BCCP, biotin carboxyl carrier protein; CT, carboxyltransferase. B, Induced and natural variants combined. Red symbols, strong alleles; blue symbols, weak or intermediate alleles. C, Highly conserved residues based on the multi-kingdom alignment of 667 sequences. D, ACC2 variants are above the horizontal bar, and ACC1 variants are below. Red bars, one accession with the predicted variant; purple bars, two to three accessions; blue bars, four to 10 accessions; green bars, more than 10 accessions; gray bars, variants were not confirmed in the only accession where it was predicted. Adapted from Parker et al. (2016).

accessions. The phenotype details of hypersensitive and sensitive accessions are defined in Chapter 3. In order to identify residues likely to be essential for protein function, we focused on ten of the most sensitive accessions from this list whose *ACC2* sequence did not contain an obvious null mutation: Sav-0, Knox-18, RRS-10, Gifu-2, Pna-10, Tul-0, Tol-0, Aitba-1, La-0, and Gn2-3. *ACC2* sequences for most of these accessions were obtained from the 1001 Genomes Project. The one exception was Sav-0, which was sequenced by Yixing Wang. Using our list of conserved residues from the 667-sequence multi-kingdom alignment, we looked for variation within these ten accessions. One group of accessions, consisting of Knox-18, RRS-10 Tul-0, and Tol-0, was predicted to contain two potential variants of interest (I404K and T1902K). Gifu-2 was predicted to have only one of those variants (T1902K). Three other sensitive accessions were predicted to have mutations affecting different conserved residues: Sav-0 (G135E), Aitba-1 (F1206L), and Pna-10 (S1883T). Neither Gn2-3 nor La-0 appeared to contain any variants in conserved amino acid residues.

Yixing Wang used Sanger sequencing to confirm these predicted variants. She found that Gifu-2 contained both I404K and T1902K variants rather than only the one predicted. Pna-10 was also found to contain the I404K and T1902K variants, and lacked the predicted S1883T mutation. Sixteen additional accessions predicted to contain these variants were tested on spectinomycin (Table 17). All of these accessions were shown to be highly sensitive to a loss of chloroplast translation, with the exception of SLSP-35 and UKSW06-333, which exhibited an intermediate phenotype. While they were predicted to have the I404K and T1902K variants, Yixing Wang confirmed that SLSP-35 and UKSW06-333 lacked both variants. The members of this group of sensitive accessions also contain another mutation affecting a conserved amino acid residue, E1355G, which can be found in a number of other accession not sensitive to spectinomycin. Results from this group of sensitive accessions provide evidence that the I404K and T1902K variants likely cause a loss of ACC2 function. Aitba-1 was confirmed to have the F1206L variant,

Table 17. Seedling Responses of 20 Accessions with the ACC2 Variants I404K and T1902K. Adapted from Parker et al. (2016).

Accession	ABRC Seed Stock	Genetic Screen	Variant Confirmation[a]	Spectinomycin Response		
				Category	Score	Seedlings
Knox-18	CS76530	Forward	A	Hypersensitive	1.1	80
RRS-10	CS76592	Forward	A	Sensitive	1.3	81
Gifu-2	CS76494	Forward	B	Sensitive	1.3	78
Tul-0	CS76618	Forward	A	Sensitive	1.7	81
Tol-0	CS76614	Forward	A	Sensitive	2.0	73
Pna-10	CS76574	Forward	C	Sensitive	1.4	112
Buckhorn Pass	CS76733	Reverse	D	Sensitive	1.7	29
Dem-4	CS76794	Reverse	D	Sensitive	1.7	28
Gre-0	CS76497	Reverse	D	Sensitive	1.4	48
MIC-31	CS77082	Reverse	D	Sensitive	2.1	49
MNF-Jac-12	CS77097	Reverse	D	Sensitive	2.5	33
MNF-Pot-21	CS77099	Reverse	D	Hypersensitive	1.4	50
MNF-Pot-75	CS77100	Reverse	D	Sensitive	2.2	32
Mdn-1	CS77077	Reverse	E	Sensitive	2.9	43
Mv-0	CS76556	Reverse	D	Hypersensitive	1.1	56
NC-6	CS77124	Reverse	D	Sensitive	1.6	54
PT2.21	CS77191	Reverse	D	Sensitive	1.5	55
Rmx-A02	CS76589	Reverse	D	Sensitive	2.1	36
Rmx-A180	CS77218	Reverse	D	Sensitive	1.8	42
SLSP-31	CS77254	Reverse	D	Sensitive	1.4	53

[a] A, Predicted sequence confirmed; B, Both variants confirmed, I404K not predicted; C, Both confirmed, neither predicted; D, Not tested; E, Both confirmed, along with E1567K (not predicted).

which is located in the middle of the central domain of ACC2.

Confirmation of the predicted amino acid variation in Sav-0 was not necessary because our lab sequenced the *ACC2* gene. Nine variable residues were found in Sav-0 compared to the consensus sequence from all Arabidopsis accessions (Table 18). Seven of these variants were in residues with low conservation in the 667-sequence multi-kingdom alignment, and were found in multiple high-intermediate or tolerant accessions. This means that these mutations are likely to have little effect on the function of ACC2 in Sav-0. One other variant (V472I) is located in a more conserved residue, but can also be found in multiple high-intermediate or tolerant accessions. On the other hand, variant G135E is located in a highly-conserved residue (95.7%), and is found only in Sav-0 and not in any other natural accession. This variant in Sav-0 is likely responsible for the hypersensitivity of the accession on spectinomycin.

Since neither Gn2-3 nor La-0 were predicted to have mutations in conserved amino acid residues, Yixing Wang sequenced the *ACC2* cDNA from both accessions to determine whether the reported sequences were correct. The La-0 sequence was identical to that reported from the 1001 Genomes Project, suggesting that the sensitivity of La-0 to a loss of chloroplast translation is caused by a defect other than a missense mutation in *ACC2*. Unlike La-0, the cDNA sequence of Gn2-3 obtained in our lab clearly differed from the 1001 Genomes sequence. However, it was identical to the sequence of Gn-1, which contains a defect in the splice acceptor site of Intron 10. Thus, the sensitivity of Gn2-3 is likely caused by this same splicing defect. From this forward genetic screen, we identified four potentially essential amino acid residues in ACC2 where missense mutations likely reduce function of the protein: G135E, I404K, F1206L, and T1902K.

Additional Accessions Chosen for Missense Mutations Affecting Conserved Residues

Using a reverse genetic approach, we identified accessions containing either single amino

Table 18. ACC2 Variants in Sav-0 that Differ from the Consensus Among Sequenced Accessions. Adapted from Parker et al. (2016).

Variant[a]	Conserv. (%)[b]	Protein Domain[c]	1001 Genomes Accessions with Predicted Variant	Accessions Evaluated on Spectinomycin[d]	Variant Confirmed	Seedlings Classified	Spectinomycin Response	
							Category	Score
A18T	Low	TP	39	Multiple[e]	Assumed	231	Intermediate	5.0
S66F	Low	TP	57	Giffo-1	Not Needed	25	High Int.	8.0
				Fell1-10; Jl-3		379	Tolerant	9.3
G135E	95.7	(BC)	0	Sav-0[f]	Yes	275	Hypersensitive	1.2
M445T	Low	BC	189	Multiple[g]	Assumed	1063	Tolerant	9.0
V472I	83.7	BC	51	Nz-1; Uk-1	Not Tested	47	High Int.	7.5
				Mt-0; Mz-0		40	Tolerant	8.8
D521N	Low	BC	51	Nz-1; Uk-1	Not Needed	47	High Int.	7.5
				Mt-0; Mz-0		40	Tolerant	8.8
S1758L	Low	CT	192	Multiple[g]	Assumed	1063	Tolerant	9.0
S2230L	Low		48	Nz-1; Uk-1	Not Needed	47	High Int.	7.5
				Mt-0		20	Tolerant	8.4
T2284R	Low		60	Nz-1; Uk-1 Lm-2; Mt-0; Mz-0	Not Needed	47	High Int.	7.5
						110	Tolerant	8.5

[a] The first residue (e.g. "G" in G135E) is found in the consensus sequence; the second in Sav-0.

[b] Conservation percentage of 667 aligned homomeric ACCase sequences with the accession consensus residue.

[c] TP, Transit Peptide; BC, Biotin carboxylase; BCCP, Biotin carboxyl carrier protein; CT, carboxyltransferase.

[d] Accessions with the same variant but a more sensitive or problematic seedling response are excluded to highlight the most tolerant responses observed with the variant present.

[e] Intermediate responses: Durh-1; Hn-0; Hovdala-2; Ler-1; Litva; Nw-0; RRS-7; Star-8.

[f] The Sav-0 variant was uncovered by sequencing the ACC2 cDNA; whole genome sequence for this accessions was not available.

[g] Tolerant responses: Fell1-10; Jl-3; Lm-2; Mt-0; Mz-0; Tsu-0; Tu-0.acid deletions, missense

mutations in the transit peptide, or missense mutations affecting conserved amino acid residues to test for sensitivity to spectinomycin. In the accession Qar-8a, there is an amino acid substitution (K1376R) immediately followed by an amino acid deletion (Δ1377). When examined on spectinomycin, the seedlings of Qar-8a were consistently sensitive, but not as highly sensitive as other accessions with null alleles of *ACC2*. This could mean that this substitution and deletion reduce but do not completely eliminate the function of ACC2. A second example of a single amino acid deletion (Δ1479) is found in the central domain of IP-Ren-6 and IP-Voz-0, but it seems to have little to no effect on the function of ACC2 since IP-Voz-0 exhibits an intermediate phenotype on spectinomycin.

Across all 855 accessions from the 1001 Genomes Project, eight variants were found in the transit peptide region of ACC2 (Table 19). We were unable to evaluate one of these (L6S) because relevant seed stocks were unavailable. Only one accession (Chi-0) contained the variant S91C. However, this variant likely has little effect on the function of ACC2 because Chi-0 has an intermediate phenotype on spectinomycin. Multiple candidate accessions were tested on spectinomycin for the other six variants. Five of these (G7V, A18T, V59L, S66F, and D87E) are found in multiple intermediate or tolerant accessions, leading us to conclude that they do not alter essential residues. Variant R4T was a promising candidate at first, based on the sensitivity of the IP-Cum-1 accession, but that variant was also confirmed in IP-Gua-1, which is an intermediate accession.

Searching the 526 conserved amino acid residues (> 95%) found in the 667-sequence multi-kingdom alignment of ACCases, we identified 44 residues where at least one Arabidopsis accession contains a missense mutation. This list of residues was evaluated further to identify essential residues where missense mutations likely reduce ACC2 protein function. Six residues were removed from the list because the accession could not be tested on spectinomycin due to lack of seeds or a known null mutation that already causes sensitivity. Four other residues were

Table 19. Accessions with Missense Mutations in the Transit Peptide of *ACC2*. Adapted from Parker et al. (2016).

Accession	ABRC Stock	ACC2 Mutation	Spectinomycin Response	
			Category	Score
Bd-0	CS76445	A-18-T & S-66-F	Low Intermediate	2.2
Bsch-0	CS76457	S-66-F	High Intermediate	7.2
Chi-0	CS76464	S-91-C	Mid Intermediate	5.1
Di-G	CS76472	A-18-T & S-66-F	Low Intermediate	2.5
Dog-4	CS76386	V-59-L	Sensitive	3.7
Fell 1-10	CS76855	S-66-F	Tolerant	8.6
Hn-0	CS76513	A-18-T & S-66-F	Mid Intermediate	5.7
IP-Cum-1	CS76787	R-4-T	Sensitive	2.3
Is-0	CS76517	S-66-F	High Intermediate	7.2
Nemrut-1	CS76398	V-59-L	Low Intermediate	4.2
Nw-0	CS76564	A-18-T & S-66-F	Mid Intermediate	7.0
RRS-7	CS76593	A-18-T & S-66-F	Mid Intermediate	4.7
Star-8	CS76400	A-18-T & S-66-F	Mid Intermediate	4.7

removed because they were not confirmed through Sanger sequencing. Appendix L lists the remaining 34 variants that alter conserved residues. From this set, 28 residues were removed from further consideration because the same variant was confirmed in at least one intermediate or tolerant accession. All six of the remaining residues are likely essential. Three of these (G135, I404, and F1206) were already identified through our forward genetic approach. One residue (E1689), with a variant found in the accession Ts-1, was already thought to be essential because it is the location of a strong mutation (*pasticcino 3-1*) in *ACC1*. The last two residues are the sites of novel missense mutations in two sensitive accessions: Y443C in Etna-2, and A2059V in Grivo-1. Through both genetic approaches to identify essential residues in ACC2, we found eight residues where a mutation likely reduces protein function of ACC2 (Table 20).

Crossing Sensitive Accessions with the Tolerant Tsu-0 Accession to Determine if *ACC2* is the Locus Responsible for Sensitivity

We took two approaches to determine whether the sensitivity of an accession with a mutation of interest was linked to the *ACC2* locus. The first approach was similar to that used to link the sensitivity of "Nossen" to *ACC2*. We crossed sensitive accessions with Tsu-0, a tolerant accession, and compared seedling phenotypes of the F2 generation to their genotypes at the *ACC2* locus. If a defect in *ACC2* was responsible for the sensitivity observed, then sensitive F2 seedlings should be homozygous for the *ACC2* allele found in the sensitive accession, whereas tolerant F2 seedlings should be homozygous or heterozygous for the Tsu-0 allele of *ACC2*. We used this method for one hypersensitive accession (Sav-0), which contains a variant (G135E) in a conserved amino acid residue, and two sensitive accessions (Nie1-2 and Oy-0), which lack an obvious defect in *ACC2*.

The results of Nie1-2 and Oy-0 crosses with Tsu-0 were at first difficult to interpret (Table 21; Figure 23). After multiple rounds of phenotyping and genotyping the *ACC2* and

Table 20. Accessions with Strong Missense Mutations Affecting Conserved ACC2 Residues. Adapted from Parker et al. (2016).

Genetic Screen	Accessions Analyzed	Variant Analyzed[a]	Protein Domain[b]	Conservation (%)[c]	Sequenced Accessions with Predicted Variant	Tolerant or Intermediate Accessions	Variant Impact[d]
Forward	Sav-0	G135E	(BC)	95.7	0[e]	0	LD
		V472I	BC	83.7	51	13	VUS
Forward	Knox-18; RRS-10; Gifu-2; Tul-0; Tol-0; Pna-10	I404K	BC	94.8	18	0	LD
		T1902K	CT-Alpha	87.6	18	0	LD
		E1355G	Central	98.7	116	18	VUS
Reverse	Etna-2	Y443C	BC	94.0	1	0	LD
Reverse	Ts-1	E1689G	CT-Beta	97.0	1	0	D

[a] The first residue (e.g. "G" in G135E) is found in the consensus sequence; the second in Sav-0.

[b] BC, Biotin carboxylase; (BC), Immediately preceding the BC domain; CT, Carboxyltransferase.

[c] Conservation percentages are based on the multi-kingdom alignment of 667 ACCase protein sequences.

[d] D, Deleterious to protein function; LD, Likely deleterious; VUS, Variant of unknown significance.

[e] Sav-0 was not included in the 1001 Genomes sequence dataset.

Table 21. Seedling Responses on Spectinomycin of Parental Accessions and F2 Progeny from Crosses between Accessions.

Genotype Examined	Total Seedlings Classified	Distribution of Seedling Phenotypes on Spectinomycin (%)[a]								
		Sensitive			Intermediate			Tolerant		
		1	2	3	5	6	7	9	10	11
Tsu-0	490		0.4	0.4	1.4	1.0	13.5	63.9	18.8	0.6
Sav-0	275	84.0	13.1	2.2		0.7				
Tsu-0 x Sav-0	428	28.5	3.0	5.9	11.9	0.7	31.1	18.0	0.9	
Oy-0	229	9.6	74.7	13.5	0.9	0.9	0.4			
Tsu-0 x Oy-0	288	0.3	7.7	14.9	30.2	2.4	27.1	14.3	3.1	
Nie1-2	235	5.5	16.6	6.0	17.9	22.5	28.1	3.4		
Tsu-0 x Nie1-2	397	0.3	0.3	0.3	5.0	1.7	23.2	48.8	17.9	2.5

[a] Numbers define classes from expanded cotyledons without leaves (1) to extensive rosettes with sizeable leaves (11) as defined in the text. Refer to Figure 3.7 for examples of seedling phenotypes for each class. Gray font, least common phenotypes (< 10%).

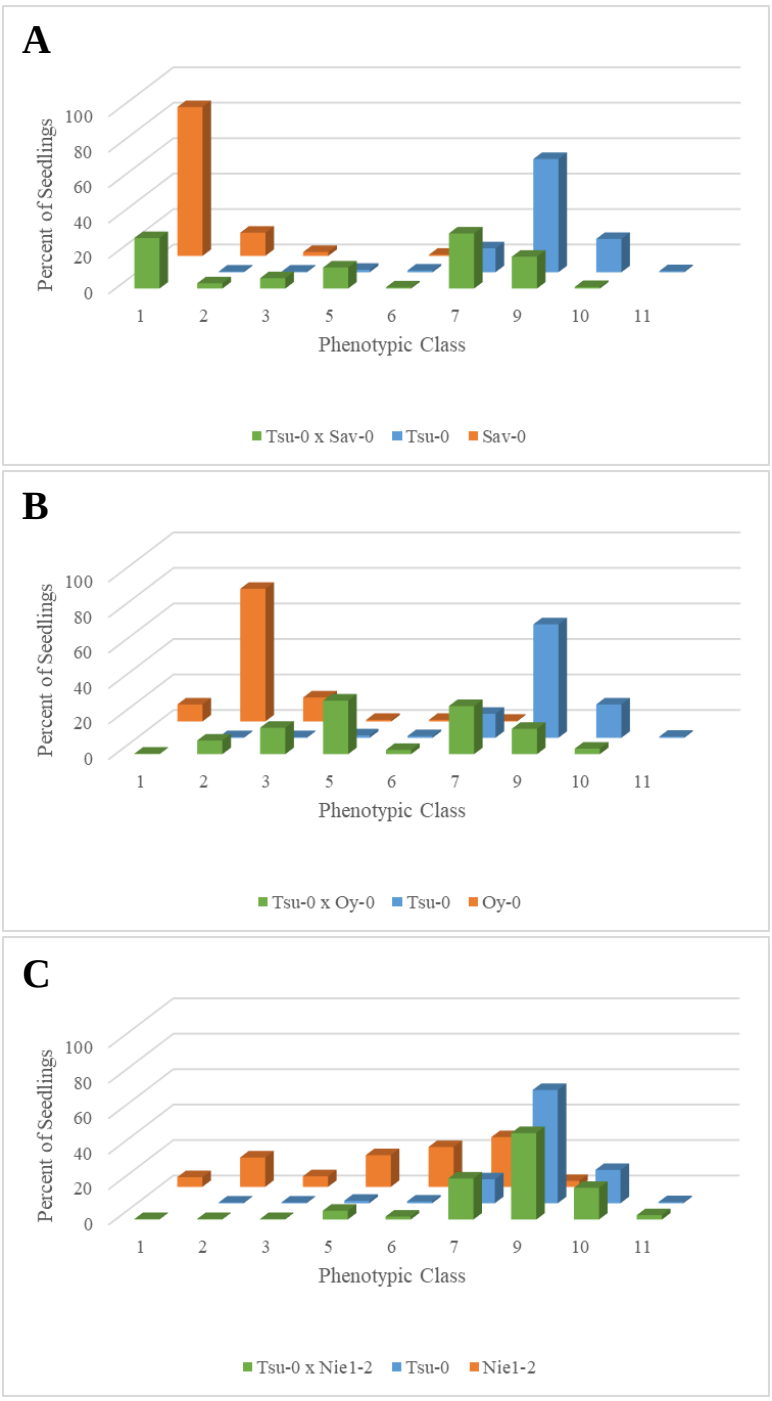


Figure 23. Comparison of Spectinomycin Seedling Responses of Parental Accessions and F2 Progeny from Crosses Between Tsu-0 and Sensitive Accessions. Percent of seedlings in each accession or F2 line assigned to the nine phenotypic categories (1-3; 5-7; 9-11) that are described in "Evaluating Additional Lines Increases the Number of Sensitive Accessions" in this Chapter. A, Tsu-0 x Sav-0; B, Tsu-0 x Oy-0; and C, Tsu-0 x Nie1-2. The data for these crosses can be found in Table 21.

enhancer loci from tolerant and sensitive F2 seedlings, we concluded that ACC2 is fully functional in Nie1-2, and that the sensitivity of this accession is due to a defect in the enhancer locus. Additional screening of Nie1-2 seedlings on spectinomycin revealed a more intermediate phenotype, which likely confirms that ACC2 is fully functional. Our current model for sensitivity in Oy-0 is a partial loss of ACC2 function, as indicated by the absences of tolerant F2 seedlings homozygous for the Oy-0 allele of *ACC2* combined with a defect in the enhancer locus. Both of these accessions show that functional ACC2 protein allows for a partial rescue of spectinomycin sensitivity, and a functional enhancer is required to increase the tolerance of an accession.

The analysis of crosses between Sav-0 and Tsu-0 revealed possible linkage between the sensitivity of Sav-0 and the genotype at the *ACC2* locus, with some inconsistent results (Table 21). While the genotype and phenotype results from the most tolerant F2 seedlings were consistent with *ACC2* as the locus responsible for Sav-0 sensitivity, results from the most sensitive F2 seedlings raised the possibility that a second locus linked to *ACC2* was responsible. However, a second round of genotyping and phenotyping of tolerant and sensitive F2 seedlings revealed perfect linkage between the sensitivity of Sav-0 and *ACC2*. These results highlighted the limitations of this approach to link sensitivity of an accession to a defect in *ACC2*. Overall, this approach proved to be a rather tedious process with results that were in some cases difficult to interpret.

Crossing Sensitive Accessions with Informative Knockout Mutants Assesses the Impact of ACC2 Variants on Protein Function

We had first assumed that our two approaches to associate sensitivity of an accession with a mutation of interest in *ACC2* would be equally informative. However, our results from crossing sensitive accessions with Tsu-0 revealed the shortcomings of that approach. Instead of continuing with those crosses, we utilized our second approach for the other sensitive accessions.

For this approach, we performed a series of genetic complementation tests using informative knockout lines of *acc2* (Salk_148966c) and *tic20-iv* (SAIL_97_F10), which are both in the Col-0 background. Both knockout lines exhibit a normal phenotype when grown in soil. However, when their seedlings are grown on spectinomycin media, they have a hypersensitive phenotype similar to accessions with null mutations in *ACC2* (Figure 24). *TIC20-IV* encodes a channel protein on the inner membrane of the chloroplast that is likely the primary channel through which ACC2, and other housekeeping proteins, pass to enter the stroma. A similar channel protein, *TOC34*, is found on the outer member of the chloroplast, and is thought to be the primary channel for movement of housekeeping proteins through the outer membrane. A *toc34* knockout mutant (*ppi3-2*) in the Col-0 background shows an intermediate phenotype on spectinomycin (Figure 24), which indicates that the Toc34 protein is not the sole channel protein for ACC2 transport through the outer membrane, whereas Tic20-IV is likely the sole channel protein through the inner membrane.

In order to analyze these genetic complementation tests, we looked at the spectinomycin phenotypes of F1 and F2 seedlings from crosses between sensitive accessions and the two knockout mutants. If a defect in *ACC2* is responsible for the sensitivity of an accession, as in the accessions with null mutations, then we expected to find 100% sensitive seedlings in both the F1 and F2 generations of the crosses with *acc2*, where the defective allele in the sensitive accession fails to complement the null allele in the knockout mutant in compound heterozygotes. For the crosses with *tic20-iv*, we also expected to see 100% intermediate seedlings in the F1 generation, where the defective alleles are complemented, and a 9:7 ratio of intermediate to sensitive seedlings in the F2 generation. The opposite is expected (100% intermediate F1 seedlings in *acc2* crosses and 100% sensitive seedlings in *tic20-iv* crosses) if the defect causing sensitivity in an accession is linked to *TIC20-IV*. If the cause of sensitivity is a defect in a gene other than *ACC2* or *TIC20-IV*, then we expected to see 100% intermediate F1 seedlings, and the 9:7 ratio of

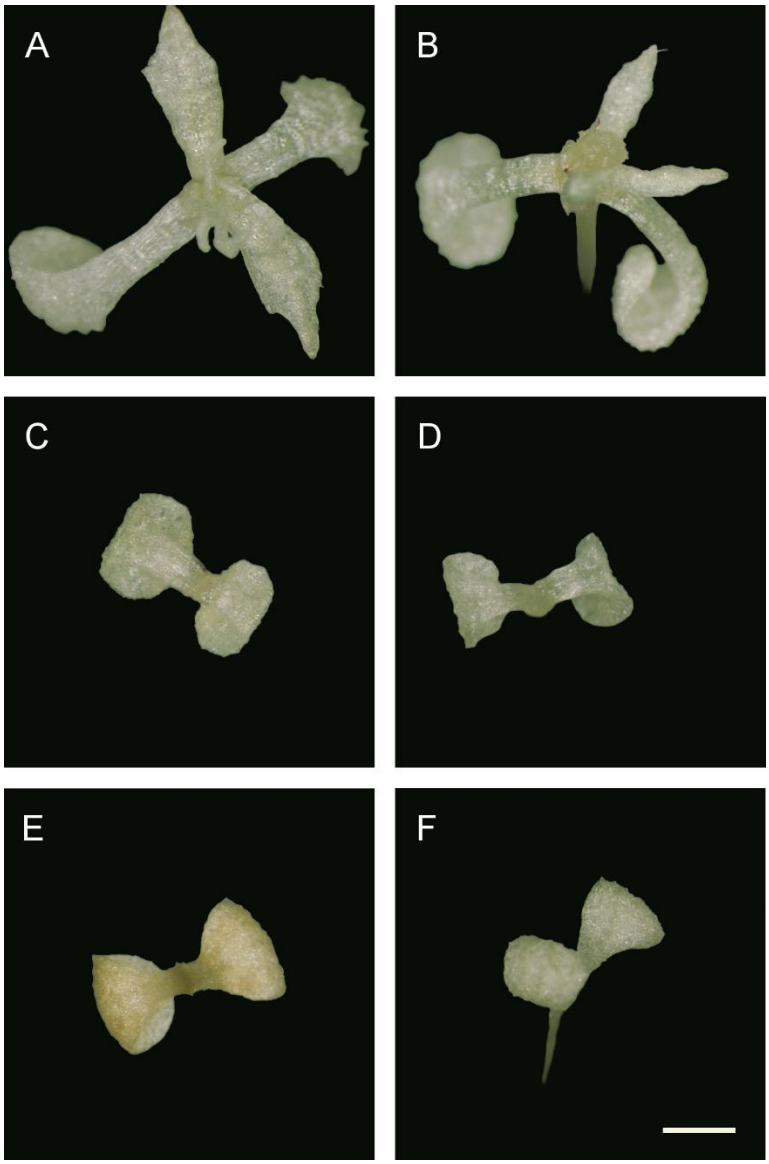


Figure 24. Spectinomycin Responses of Knockout Mutants of Known Components of the Chloroplast Protein Import System. A, Parental Col-0 accession. B, *toc34-1* (*ppi3-2*). C, *tic20-iv-1* (SAIL_97_F10). D, *tic20-iv-2* (Koncz 11324). E, *acc2-1* (Salk_148966c). F, Sav-0 (the most sensitive accession). Bar = 1 mm. Adapted from Parker et al. (2014).

intermediate to sensitive F2 seedlings.

This approach was used for a total of 17 sensitive accessions: Gn2-3 as a control; Sav-0; Knox-18, Gifu-2, Pna-10, RRS-10, and Tul-0 as representatives of the large group with two missense mutations (I404K and T1902K); five additional accessions with different missense mutations thought to reduce the function of ACC2 (Aitba-1, F1206L; Etna-2, Y443C; Grivo-1, A2059V; Ts-1, E1689G; and Qar-8a, K1376R and Δ1377); two accessions where there is no obvious defect in *ACC2* (IP-Cum-1 and La-0); and three accessions with missense mutations of unknown significance in *TIC20-IV* (IP-Deh-1, IP-Tdc-0, and Kru-3). The results of these crosses are shown in Tables 22 and 23 along with Figure 25.

Previously, we found that Gn2-3 has a splicing defect in the acceptor site of Intron 10, which causes a frameshift and a truncated ACC2 protein. We used Gn2-3 as a control to see if our predictions were correct for accessions whose sensitivity is caused by a defect in *ACC2*. The results of these crosses showed exactly what we expected. For the crosses with *acc2*, all of the seedlings from the F1 and F2 generations were sensitive to spectinomycin, with a sensitive phenotype similar to both parent lines. The F1 seedlings from the *tic20-iv* crosses were all less sensitive than the parent lines, and a majority of them had an intermediate phenotype. In the F2 generation, 56% of the seedlings showed an intermediate phenotype while 44% were sensitive, which almost perfectly matches our expected 9:7 ratio. The results of the Sav-0 crosses followed the same pattern. All of the seedlings in the F1 and F2 generation of the *acc2* crosses were sensitive to spectinomycin while the F1 seedlings from the *tic20-iv* crosses were all intermediate and the F2 seedlings were 61% intermediate to 39% sensitive, which is still close to the expected 9:7 ratio. This provides substantially more evidence that the sensitivity of Sav-0 is connected to the G135E substitution in ACC2. Due to the hypersensitivity of Sav-0 seedlings, this missense mutation likely eliminates ACC2 protein function.

All five accessions (Knox-18, Gifu-2, Pna-10, RRS-10, and Tul-0) used to represent the

Table 22. Spectinomycin Responses of F1 Seedlings from Crosses between Sensitive Accessions and Informative Knockout Mutants. Adapted from Parker et al. (2016).

Accession Parent	F1 Progeny from *acc2* Cross			F1 Progeny from *tic20-iv* Cross		
	Category	Score	Seedlings	Category	Score	Seedlings
"Gn2-3"	Hypersensitive	1.1	106	Intermediate	4.9	98
Sav-0	Hypersensitive	1.0	71	Intermediate	4.4	45
Knox-18[a]	Sensitive	1.3	78	Intermediate	7.2	83
Gifu-2[a]	Sensitive	1.2	82	Intermediate	5.9	72
Pna-10[a]	Hypersensitive	1.0	74	Intermediate	5.0	64
RRS-10[a]	Sensitive	1.2	80	Intermediate	5.8	81
Tul-0[a]	Hypersensitive	1.1	83	Intermediate	5.8	81
Aitba-1	Sensitive	2.1	149	Intermediate	5.4	148
Etna-2	Intermediate	5.8	89	Intermediate	5.6	100
Grivo-1	Intermediate	7.3	74	Intermediate	6.6	52
Ts-1	Sensitive	1.7	106	Intermediate	5.6	81
Qar-8a	Low Intermediate	3.6	96	Intermediate	4.9	107
IP-Cum-1	Intermediate	4.9	83	Intermediate	4.4	62
La-0	Intermediate	4.6	61	Intermediate	4.6	67
IP-Deh-1[b]	Intermediate	6.9	56	Intermediate	3.9	81
IP-Tdc-0[b]	Intermediate	7.1	86	Intermediate	6.7	87
Kru-3[b]	Intermediate	7.1	80	Intermediate	5.7	87

[a] Part of the Knox-18 group of sensitive accessions with shared variants of interest.

[b] Contains a missense mutation of unknown significance in *TIC20-IV*.

Table 23. Spectinomycin Responses of F2 Seedlings from Crosses between Sensitive Accessions and Informative Knockout Mutants. Adapted from Parker et al. (2016).

Accession Parent	Knockout Parent	Seedlings Classified	Phenotype Score	Distribution of F2 Seedling Phenotypes (%)							
				Sensitive			Intermediate			Tolerant	
				1	2	3	5	6	7	9	10
"Gn2-3"	*acc2*	82	1.0	97.6	2.4						
Sav-0	*acc2*	163	1.0	99.4	0.6						
Sav-0	"Nossen"	184	1.2	87.5	4.9	7.1		0.5			
Knox-18	*acc2*	127	1.0	96.1	3.9						
Tul-0	*acc2*	135	1.1	90.4	8.9	0.7					
Aitba-1	*acc2*	325	2.2	42.5	28.6	17.2	5.5	3.1	3.1		
"Gn2-3"	*tic20-iv*	156	3.6	39.1	5.1	5.1	30.8	1.9	16.7	1.3	
Sav-0	*tic20-iv*	152	3.2	31.6	7.9	25.7	26.3	0.7	7.8		
Knox-18	*tic20-iv*	124	4.0	46.0	4.8		12.1	4.0	21.8	10.5	0.8
Tul-0	*tic20-iv*	138	3.8	38.4	10.9		15.2	10.1	21.0	4.4	
Aitba-1	*tic20-iv*	124	3.6	29.9	11.3	15.3	21.0	4.8	13.7	4.0	
La-0	*acc2*	152	2.6	23.7	27.6	35.5	5.9	6.6	0.7		
La-0	*tic20-iv*	154	2.5	25.3	29.2	35.1	5.9	4.5			
Ip-Cum-1	*acc2*	112	4.4	16.1	9.8	11.6	31.3	8.9	22.3		
Ip-Cum-1	Sav-0	199	3.1	30.7	9.0	27.1	26.1	1.1	6.0		
Ip-Cum-1	*tic20-iv*	229	3.5	22.3	10.0	24.0	35.8	1.8	6.1		

[a] Numbers define classes from expanded cotyledons without leaves (1) to extensive rosettes with sizeable leaves (11) as defined in the text. Refer to Figure 3.7 for examples of seedling phenotypes for each class. Red font, most common phenotypes (> 10%). Gray font, least common phenotypes (< 5%).

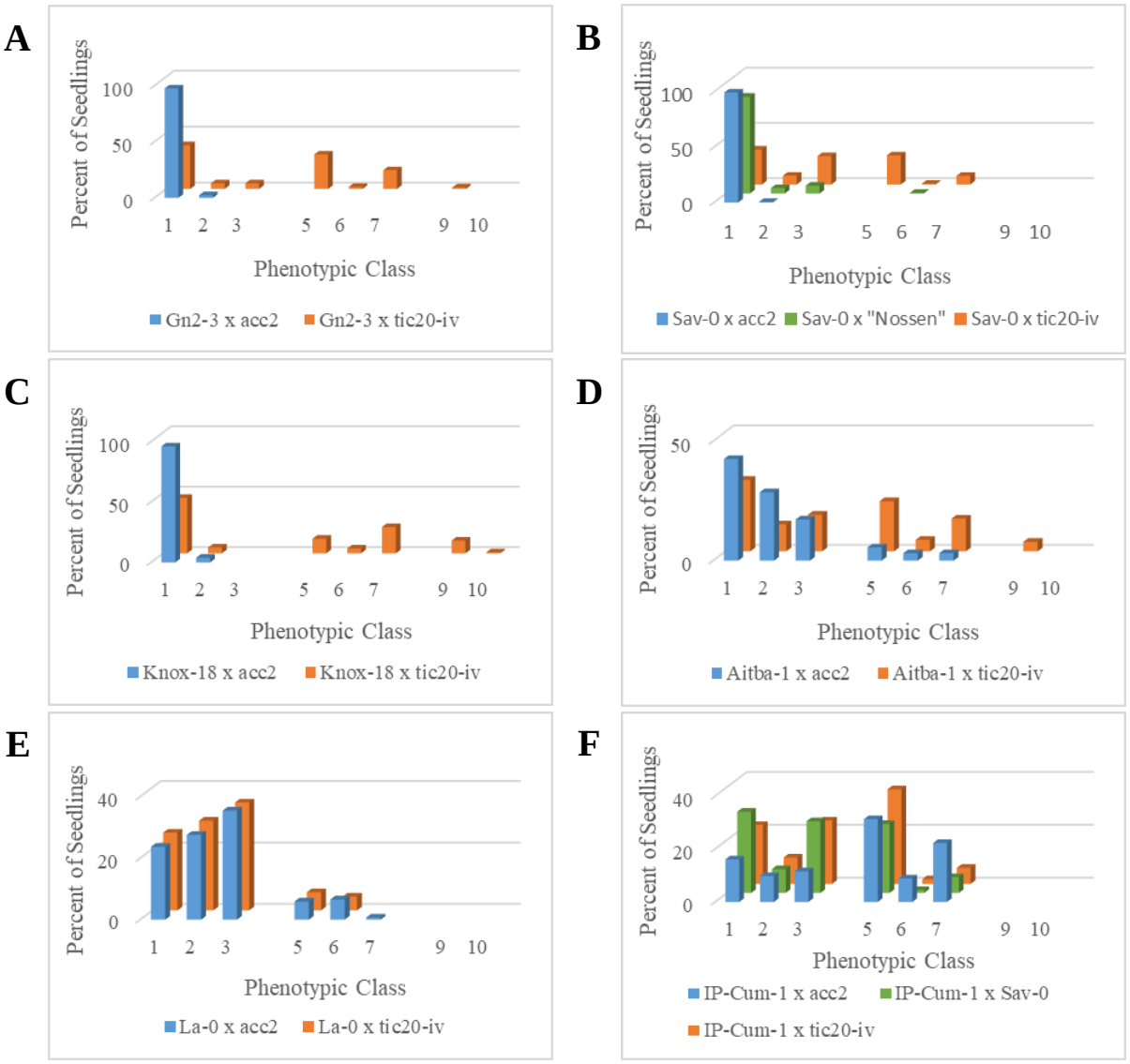


Figure 22. Comparison of Spectinomycin Seedling Responses of F2 Seedlings from Crosses between Sensitive Accessions and Informative Knockout Mutants. Percent of F2 seedlings in each cross assigned to the nine phenotypic categories (1-3; 5-7; 9-10) that are described in "Evaluating Additional Lines Increases the Number of Sensitive Accessions" in this Chapter. No seedlings were found in category 11. A, Gn2-3 crossed with *acc2* and *tic20-iv*; B, Sav-0 crossed with *acc2*, "Nossen" and *tic20-iv*; C, Knox-18 crossed with *acc2* and *tic20-iv* (Tul-0 crosses showed a similar graph); D, Aitba-1 crossed with *acc2* and *tic20-iv*; E, La-0 crossed with *acc2* and *tic20-iv*; and F, IP-Cum-1 crossed with *acc2*, Sav-0 and *tic20-iv*. The data for these crosses can be found in Table 23.

group of 20 accessions with the I404K and T1902K variants, showed results similar to the Sav-0 and Gn2-3 crosses. Knox-18 and Tul-0 were taken to the F2 generation, whereas Gifu-2, Pna-10, and RRS-10 were only analyzed at the F1 generation. In all five cases, all of the seedlings from the crosses with *acc2* were sensitive to spectinomycin, with the exception of two seedlings from the Knox-18 cross that were intermediate. The majority of the F1 seedlings from all five *tic20-iv* crosses had an intermediate phenotype, and the F2 seedlings from the Knox-18 and Tul-0 crosses showed about a 1:1 ratio of intermediate to sensitive seedlings, which could resolve into a 9:7 ratio if more seedlings were screened. These results indicate that the two *ACC2* missense mutations in this group of accessions are responsible for the sensitive phenotype of all 20 members of the group. In order to be sure that some other null mutation is not present in these accessions, Yixing Wang sequenced the *ACC2* cDNA from Knox-18, and confirmed that a full length transcript is produced.

Of the crosses with the other five accessions that contain missense mutations that we originally thought affected the function of ACC2, only the crosses with Ts-1 showed *ACC2* as the locus responsible for sensitivity of the accession. The F1 seedlings from the Ts-1 crosses with *acc2* were sensitive to spectinomycin while the F1 seedlings from the crosses with *tic20-iv* were intermediate. This result is not surprising since the position of the missense mutation in Ts-1 (1689) is the same as a strong mutation (*pasticcino 3-1*) in *ACC1*. In Ts-1, the missense mutation, E1689G, likely reduces ACC2 function significantly. The results of the Aitba-1 crosses may indicate that the missense mutation (F1206L) in *ACC2* is responsible for sensitivity, but they are harder to interpret than the other crosses. This is likely due to Aitba-1 seedlings having a less sensitive phenotype than accessions with null mutations in *ACC2*. In the crosses with *acc2*, the F1 seedlings were all sensitive to spectinomycin, similar to the parent lines. A majority of the F2 seedlings were sensitive, but around 11% of the seedlings were intermediate. About 75% of the F1 seedlings from the Aitba-1 crosses with *tic20-iv* were intermediate while the rest were

sensitive. The results of the F2 generation are harder to interpret because it is difficult to distinguish the phenotypes of high sensitive and low intermediate seedlings. The F1206L missense mutation in Aitba-1 likely reduces the function of ACC2, but not as severely as a null mutation.

Results from the Grivo-1 crosses clearly showed that the sensitivity of Grivo-1 is not associated with either *ACC2* or *TIC20-IV*. The F1 seedlings from both crosses were all intermediate when compared to the phenotypes of the parent lines. The F2 generation was not studied since harvesting F2 seeds would have required a vernalization treatment of 5-6 weeks. Results from the Etna-2 and Qar-8a crosses showed that the sensitivity of these accessions is also likely caused by a defect in a gene other than *ACC2* or *TIC20-IV*, though the results are less definitive than Grivo-1. In both cases, F1 seedlings from the *tic20-iv* crosses were almost all intermediate, whereas around 20% of the F1 seedlings from the *acc2* crosses were sensitive. Again, the F2 generation was not analyzed due to the requirements of harvesting F2 seed. Overall, the missense mutations in Grivo-1 (A2059V), Etna-2 (Y443C), and Qar-8a (K1376R and Δ1377) are examples of substitutions in highly conserved residues that do not appear to reduce the function of ACC2.

Results from crosses with two accessions (IP-Cum-1 and La-0) that lack obvious defects in either *ACC2* or *TIC20-IV*, and three accessions (IP-Deh-1, IP-Tdc-0, and Kru-3) with missense mutations in *TIC20-IV*, showed that the cause of sensitivity in these accessions was not associated with either locus. The F1 seedlings produced when these accessions were crossed with *acc2* and *tic20-iv* all showed an intermediate phenotype more tolerant than any of the parent lines on spectinomycin. The F2 generation of the IP-Cum-1 and La-0 crosses were also analyzed, and the results did not appear to show the 9:7 ratio we expected to see if a single locus was responsible for the sensitivity of the accessions.

Through all of these crosses, we have shown that a single missense mutation can cause a

partial or full loss of ACC2 function in Sav-0 (G135E), the group of 20 accessions (I404K and T1902K), Aitba-1 (F1206L), and Ts-1 (E1689G). We have also shown that sensitivity of some accessions can be linked to the enhancer locus on chromosome 5, as in Oy-0 and Nie1-2, and we have eight sensitive accessions where the defect responsible for sensitivity is not located in *ACC2* or *TIC20-IV*.

Tolerated Missense Mutations in *ACC1* and *ACC2*

So far, we have been using accessions sensitive to spectinomycin to look for informative missense mutations in ACC2 that severely reduce or eliminate function of the protein. However, this represents a small fraction of the total variation found in ACCase protein sequences among natural accessions. In order to look at natural variation in the paralogous ACC1 protein sequences, we utilized an Excel spreadsheet similar to the one described for comparing ACC2 sequences. Any missense mutations found in *ACC1* must be tolerated as it is an essential protein, and the loss of ACC1 function results in seedling lethality. We identified 132 variable residues in ACC1 across all sequenced accessions. Nineteen residues were removed from this list because their variation was likely due to errors in sequencing rather than a true substitution as indicated by the presence of an unknown amino acid ("Z") in the protein sequence and unresolved nucleotides in the genomic sequence. Appendix I lists the 113 variable residues in the ACC1 protein sequence found among 855 natural Arabidopsis accessions. These 113 residues are spread throughout the protein sequence, with the highest concentration (33%) located in the central domain of the protein, where the most variation was also seen among the ACC2 protein sequences. Seven of these predicted variable residues are highly conserved throughout our multi-kingdom alignment of 667 sequences. Using Sanger sequencing, Yixing Wang confirmed two of the variants in conserved residues, A193V and V809A, and did not confirm the presence of three other variants

(A271D, Q272R and L742S) in the only accession where they were predicted. In total, we found 110 variable residues in the ACC1 protein sequence among 855 natural Arabidopsis accessions, which is only 5% of the entire ACC1 sequence.

In addition to the natural variation found in ACC1 protein sequences, we identified tolerated missense mutations using the ACC2 protein sequences from tolerant accessions, which likely have a fully functional ACC2 protein. We aligned the protein sequences from eight of the most tolerant accessions: Chat-1, Ema-1, Ha-HBT1-2, Lm-2, Pog-0, Tsu-0, Tu-0, and Uod-1. Using this alignment, we found 24 total residues where substitutions are tolerated. None of these variants were found in all tolerant accessions, indicating that a single missense mutation is likely not responsible for spectinomycin tolerance and the consensus sequence from all 857 accessions encodes a fully functional ACCase protein. Similar to the variation found among the ACC1 and ACC2 protein sequences for all accessions, most (42%) of the variation found in the ACC2 sequences of tolerant accessions is within the central domain of the protein. Information about these variable residues in tolerant accessions is listed in Table 24. Remarkably, six of the 24 variants (P475L, Q478K, N725S, R762C, E1355G, and G1766D) are in highly conserved residues found through our multi-kingdom alignment of 667 sequences. In addition, from our crosses with *acc2* and *tic20-iv* knockout mutants, we have confirmed three missense mutations that likely do not reduce ACC2 protein function.

We also found variation in 18 other highly conserved residues where the accessions associated with the variant showed an intermediate phenotype on spectinomycin, indicating that these missense mutations may slightly reduce ACC2 protein function. Table 25 lists the 24 variants found in highly conserved residues where there is evidence of at least partial ACC2 protein function. Overall, we found 137 residues (6% of the total residues) in ACCase protein sequences that can tolerate missense mutations without affecting protein function, 13 of which are located in residues highly conserved among 667 eukaryotic ACCase sequences. Additionally, we

Table 24. ACC2 Variants Found in the Most Tolerant Natural Accessions. Adapted from Parker et al. (2016).

Variant Analyzed [a]	Protein Domain [b]	Conservation (%) [c]	Sequenced Accessions with Predicted Variant	Tolerant Accessions
G7V	TP	Low	175	Ema-1; Ha-HBT1-2
D87E	TP	Low	333	Chat-1; Ema-1; Pog-0
D101G	-	Low	82	Ema-1
A132S	-	27.9	82	Ema-1
G355V	BC	30.1	130	Ha-HBT1-2; Uod-1
M445T	BC	Low	190	Lm-2; Tsu-0; Tu-0
P475L	BC	99.7	1	Lm-2
Q478K	BC	97.6	28	Uod-1
N725S	-	96.9	44	Chat-1; Pog-0
R762C	-	96.6	6	Tsu-0; Tu-0
Q903H	Central	29.5	78	Ema-1
S949F	Central	Low	109	Ema-1
E975K	Central	Low	192	Ha-HBT1-2; Uod-1
E1103K	Central	Low	210	Ha-HBT1-2; Uod-1
T1238I	Central	Low	10	Ha-HBT1-2
E1312D	Central	Low	116	Ema-1
E1355G	Central	98.7	116	Ema-1
T1384S	Central	Low	34	Chat-1; Pog-0
I1403N	Central	Low	79	Ema-1
G1420A	Central	12.3	7	Tsu-0; Tu-0
S1758L	CT-β	Low	193	Lm-2; Tsu-0; Tu-0
G1766D	CT-β	97.6	34	Chat-1; Pog-0
N1961D	CT-α	Low	7	Uod-1
T2284R	-	Low	61	Lm-2

[a] The first residue (e.g. "G" in G7V) is found in the consensus sequence; the second in Ema-1.

[b] TP, Transit peptide; BC, Biotin carboxylase; CT, Carboxyltransferase.

[c] Conservation percentages are based on the multi-kingdom alignment of 667 ACCase protein sequences.

Table 25. Accessions with Evidence of Residual ACC2 Function Despite Substitutions in Highly-Conserved Residues. Adapted from Parker et al. (2016).

Variant Analyzed[a]	Protein Domain[b]	Conservation (%)[c]	Sequenced Accessions with Predicted Variant	Intermediate; Low-Intermediate Accessions[d]	Tolerant; High-Intermediate Accessions[e]
F363L	BC	99.3	5	Sei-0	
V376A	BC	100.0	12	Col-0	
L474F	BC	94.5	1	Chi-0	
P475L	BC	99.7	1		Lm-2
Q478K	BC	97.6	28	Multiple	Uod-1
R494G	BC	99.9	1	Ip-Pal-0	
T538A	BC	99.9	1	IP-Tor-1	
N725S	-	96.9	44	Multiple	Pog-0
G739E	-	95.2	1	Wa-1	
R762C	-	96.6	6	Mh-0	Tsu-0; Tu-0
G833R	BCCP	99.3	3	Dja-1	
L847P	-	96.0	1	WAR	
E1355G	Central	98.7	116	Multiple	Si-0; Ema-1
R1405Q	-	96.1	1	Db-1	
G1766D	CT-Beta	97.6	39	Multiple	Pog-0
I1821V	CT-Beta	98.2	1	MNF-Che-2	
T1834S	CT-Beta	99.4	2	Nemrut-1	
S1883T	CT-Beta	97.0	8	Multiple	
G1897S	CT-Alpha	99.4	2	Sch1-7; WalHaesB4	
P2013L	CT-Alpha	98.5	3	Balan-1	
A2014E	CT-Alpha	99.0	1	App1-16	
I2115R	CT-Alpha	98.2	1	Iasi-1	
H2207Q	CT-Alpha	98.1	1	Ip-Lso-0	

[a] The first residue (e.g. "F" in F363L) is found in the consensus sequence; the second in Sei-0.

[b] BC, Biotin carboxylase; BCCP, Biotin Carboxyl Carrier Protein; CT, Carboxyltransferase.

[c] Conservation percentages are based on the multi-kingdom alignment of 667 ACCase protein sequences.

[d] May contain partial loss-of-function alleles of *ACC2*.

[e] Likely contain fully-functional alleles of *ACC2*.

found 18 highly conserved residues in ACC2 where it seems that missense mutations lead to a partial loss of protein function.

DISCUSSION

This chapter describes the use of forward and reverse genetic approaches to identify residues in *ACC2* that are likely essential for full protein function. This experimental system using spectinomycin to evaluate the level of function of ACC2 in natural accessions of Arabidopsis is a unique way to analyze the effects of mutations on a highly-conserved, essential gene in fatty acid biosynthesis. Two advantages of this system are: (1) while ACC1 plays a key role in Arabidopsis growth and development, ACC2 is essential only when chloroplast translation, and consequently the production of the heteromeric ACCase, is blocked; and (2) utilizing spectinomycin to inhibit chloroplast translation provides a method to analyze the effects of mutations on ACC2 at the seedling level. Null mutations in other ACCase proteins lead to lethality, but null mutations in *ACC2* result in an easy-to-identify hypersensitive phenotype on spectinomycin. Prior to this study, relatively few studies had been published on missense mutations in ACCase proteins. Arabidopsis was a key player in these studies because the mutations can be maintained as heterozygotes and the effects studied in segregating seeds and embryos (Meinke et al., 2008). Both strong and weak mutant alleles of *ACC1* have been used to understand the function of ACCases in Arabidopsis (Meinke, 1985; Baud et al., 2004; Kajiwara et al., 2004; Lu et al., 2011; Amid et al., 2012). While other ACCase mutations are found in *Caenorhabditis elegans* (Rappleye et al., 2003), *Drosophila melanogaster* (Sasmura et al., 2013), and *Saccharomyces cerevisiae* (Schneiter et al., 1996, 2000), most of these mutations offer little evidence on ACCase protein function. Several of the missense mutations in *S. cerevisiae* provide some information on key regions in the dimer interface (Wei and Tong, 2015). More recently, the focus of ACCase research has been on identifying herbicide resistant mutations in grasses

(Kaundun, 2014) and using ACCases as targets for drugs such as antibiotics, antifungals, and those for obesity and type-2 diabetes (Campbell and Cronan, 2001; Lenhard, 2011; Tong, 2013). Prior to our work, fewer than 20 amino acid residues in ACCase proteins had been associated with mutations affecting protein function. We have expanded this list of residues using the mutations found in natural accessions of Arabidopsis. The mutations from these previous studies, along with those found in this study, are listed in Appendix K.

The K_a/K_s analysis of the Brassicaceae ACCase sequences raises a question about the function of *ACC2* in natural accessions of Arabidopsis: Why is there evidence of purifying selection on *ACC2* when the gene is not essential for survival? One possible answer is that ACC2 has a function outside of its known involvement in the conversion of malonyl-CoA to acetyl-CoA in the chloroplast. Potentially, ACC2 could function in a metabolic pathway within the mitochondria. This would be similar to the duplicated ACCase found in mammals and *S. cerevisiae*, which has been shown to function in the oxidation of fatty acids (Hoja et al., 2004; Abu-Elheiga et al., 2005). If ACC2 functions in a mitochondrial metabolic pathway, the loss of ACC2 function in some natural Arabidopsis accessions might indicate that the pathway is either not crucial for the plant's survival or the loss of ACC2's function in the pathway can be compensated by another protein. A second possible reason that *ACC2* has remained functional in most accessions is that ACC2 converts acetyl-CoA to malonyl-CoA in the chloroplast when the heteromeric ACCase protein is post-translationally down-regulated by the buildup of fatty acids in the endoplasmic reticulum (ER; Bates et al., 2014). In this case, there may be some advantages for Arabidopsis plants in selected environments to continue synthesizing fatty acids even if they accumulate in the cell. However, it is unlikely that continued synthesis of fatty acids is required for plant growth and development, which would explain why some natural accessions lack ACC2 function. Expanding the K_a/K_s analysis using additional ACCase sequences from other members of the Brassicaceae might help to resolve these questions.

In our study of ACC2 protein sequences from the 1001 Genomes Project, we identified four nonsense mutations, two single-nucleotide deletions that caused frameshifts, three defects in RNA splicing, two large deletions or chromosomal rearrangements, one small deletion that caused a frameshift, and five essential amino acid residues where missense mutations or single-amino acid deletions are found in natural Arabidopsis accessions. All of these mutations likely have effects on the structure and function of ACC2. Other potential defects in *ACC2* affecting protein function that we have not identified include mutations in the promoter region, which are difficult to evaluate based on sequence variation alone, changes in the 5' or 3' untranslated region of the mRNA, which could affect the initiation or termination of translation, or mutations that reduce translation efficiency. Yixing Wang tested for promoter defects that reduce the amount of *ACC2* transcript in a small number of sensitive accessions, including "Nossen", Nie1-2, and Oy-0, using qRT-PCR experiments, which showed no difference between the amount of *ACC2* transcript produced from these sensitive accessions and multiple tolerant accessions. However, this does not rule out the possibility of a promoter defect in other sensitive accessions.

Similar to the approach used in human genetics to describe missense mutations that cause a phenotype, we divided the variants found in our study into six categories based on the effects of the variant on ACCase protein function: deleterious, likely deleterious, potentially deleterious, variant of unknown significances, likely not deleterious, and not deleterious (Parker et al., 2016). Through our genetic complementation tests, we confirmed five variants that significantly reduced or completely eliminated ACC2 protein function: G135E, I404K, F1206L, E1689G, and T1902K. The reduction of ACC2 function caused by F1206L is likely not as severe as the other four missense mutations, because Aitba-1 is one of the less sensitive accessions. Four of the five missense mutations (G135E, I404K, F1206L, and T1902K) shown through genetic complementation tests to impact the structure and function of ACC2 are categorized as likely deleterious while the fifth mutation (E1689G) was labeled as deleterious. These genetic

complementation tests also revealed three missense mutations in highly conserved residues (Y443C, K1376R, and A2059V) that seem to have no effect on ACC2 function. These substitutions are categorized as likely not deleterious.

The 110 variants found in the comparison of ACC1 protein sequences among natural Arabidopsis accessions were labeled as likely not deleterious substitutions when the variant is predicted in a single accession, or not deleterious substitutions when the variant is predicted in more than one accession. Of the 24 missense mutations found in the ACC2 protein sequences of tolerant accessions, 18 are categorized as variants of unknown significance because they are found in at least one tolerant accession, but no other information is known about those residues. Five of the other missense mutations are categorized as likely not deleterious because they are found in tolerant accessions and are located in highly conserved amino acid residues. The last missense mutation, found in Tsu-0 and Tu-0, is categorized as not deleterious since it is located in a highly conserved residue, and we have substantial evidence that the Tsu-0 allele of ACC2 is fully functional. Additionally, there are 179 variants categorized as likely not deleterious because the consensus sequences of ACC1 and ACC2 from the natural accessions differ, and both consensus sequences encode functional ACCase proteins. Any variation found in the ACC1 protein sequences of natural accessions cannot be considered deleterious because *ACC1* is an essential gene.

Through this study, we have identified 18 missense mutations that slightly reduce the function of ACC2. These 18 mutations are located in highly conserved residues. The furthest the seedlings from any accession with one of these residues develop is to an intermediate stage. Of these 18 variants, V376A is the most interesting. This mutation is located in an amino acid residue that is perfectly conserved in our multi-kingdom alignment of 667 sequences, and among all natural accessions of Arabidopsis, with the exception of the Col-0 protein sequence, which is

the reference Arabidopsis accession. The conservation of this residue, and the mutation only found in Col-0, raises the possibility that the Col-0 ACC2 protein has reduced function.

With our combined approaches, we analyzed the effects of 339 different missense mutations on ACCase function. However, this represents only 15% of the total residues found in an ACCase protein, which leaves around 85% of the residues to be analyzed in order to fully understand the effects of missense mutations on ACCase function. This highlights the limitations of utilizing natural variation to study the effects of mutations on protein structure and function. In order to learn more about missense mutations in ACCase proteins, this project would need to be expanded using recent advances in gene editing technologies to induce missense mutations in residues of interest. For example, the A376V mutation found in Col-0 could be induced in the ACC2 sequence of Tsu-0, or another tolerant accession, which can then be analyzed on spectinomycin to look for a reduction of ACC2 function indicated by increased sensitivity. In order to evaluate some of the more subtle changes in ACCase protein function, missense mutations could be induced in candidate residues within ACC1, where the Meinke lab has shown, using *emb22*, *pas3-1*, and *pas3-2* mutants, that the strength of the mutation affects the terminal embryo phenotype (Parker et al., 2016).

CHAPTER VI

CANDIDATE GENE APPROACH TO IDENTIFY OTHER FACTORS THAT INCREASE TOLERANCE TO A LOSS OF CHLOROPLAST TRANSLATION

INTRODUCTION

As described in Chapter 4, we used crosses between *emb* mutants defective in chloroplast translation and the tolerant Tsu-0 accession to identify a single, dominant suppressor that increases tolerance of a loss of chloroplast translation. We also found evidence for a second, unlinked locus that enhances the effect of the suppressor, and additional genetic modifiers that further increase tolerance. Through our analysis of these crosses, we identified the suppressor locus as *ACC2* and mapped the enhancer near the top of chromosome 5 based on tight linkage with *EMB3137*. Further work on identifying the enhancer was performed by Kayla Cook, an undergraduate researcher in our lab (Cook and Meinke, 2017). Kayla manually curated the region of chromosome 5 surrounding *EMB3137* to identify potential candidates for the enhancer locus. Through this curation, she found seven candidate genes that encode proteins with functions consistent with one of our models for the enhancer.

For this final part of my project, I focused on a method to identify additional genetic

modifiers of this system. Unlike the evidence we have for a single suppressor locus and enhancer, there seem to be at least two modifier loci that have some effect even without a functional enhancer. However, the modifiers appear to require the presence of the enhancer to have a significant impact on tolerance. In order to narrow the search for these modifiers, I used a candidate gene approach focused on five components of the TIC/TOC chloroplast protein import system that are found in different regions of the Arabidopsis genome, which allowed us to examine the regions surrounding these loci for linkage between a candidate gene and a potential modifier. Five descendent lines were used to compare the genotype of each potential modifier to the differences in embryo rescue. Two groups of lines were used: those likely to be homozygous Tsu-0 for each of the modifiers, and those likely to be homozygous "Nossen". Unfortunately, no association was found between the genotype of each candidate gene and the amount of embryo rescue observed in the descendent lines. Candidate modifiers that have not been tested yet include four additional members of the TIC/TOC system and ten gene products that likely interact with one or more of the potential enhancers identified through Kayla's curation. In contrast to previous sections of this dissertation, none of the work described in this chapter has been published.

MATERIALS AND METHODS

Plant Material and Growth Conditions

Mature seeds from the F4 and F5 generations of a cross between the tolerant Tsu-0 accession and *emb3126-1*, a mutant defective in chloroplast translation in the sensitive "Nossen" background, were harvested in our laboratory. These seeds were then germinated on plates containing a basal nutrient medium as described in Chapter 3. After plating the seeds, the plates were stored at 4° C in a refrigerator for three days, and then placed under fluorescent lights for 14 days at room temperature. Seedlings were then divided into two groups: some were used for DNA

extractions (described later), and others were transplanted to pots and grown in a growth room as described in Chapter 3. After four weeks, seed and embryo measurements were taken using the method described in Chapter 4.

PCR Genotyping of Plants

For each descendent line used in this analysis, genomic DNA was extracted from six seedlings for PCR genotyping. Additional seedlings were frozen as backups if more DNA was needed. Genomic DNA extraction was performed using a modified cetyltrimethylammonium bromide protocol (Lukowitz et al., 2000). Following the DNA extraction, specific loci (described later) were amplified through PCR using the Qiagen PCR Master Mix and a Biometra Uno II thermocycler. The PCR primers used for each locus were designed by Yixing Wang based on polymorphic differences between the Tsu-0 and "Nossen" genomic sequences (Table 26). Tsu-0 sequences were obtained through the 1001 Genomes Project database, whereas "Nossen" sequences were obtained from the laboratory of Dr. Masatomo Kobayashi at the RIKEN BioResource Center. All primers were purchased from Integrated DNA Technologies. PCR products were separated in 1% agarose gels containing GelRed Nucleic Acid Stain (Phenix), and bands were visualized using the AlphaImager HP system (Proteinsimple). These products were then purified using the QIAquick PCR purification kit (Qiagen), and sent for sequencing at the Oklahoma State University Recombinant DNA/Protein Resource Facility. Sequencing results were visualized for analysis using FinchTV version 1.4.0 (Geospiza Inc.).

Loci Chosen as Modifier Candidates

For this candidate gene approach to identify potential modifiers, we chose five loci, each

Table 26. PCR Primer Sequences Used for Plant Genotyping.

Name	Primer Sequence	Primer Location
1-2-F 205	CCTGCATCAATGAAGGGATTTG	Intron 2 of *TIC110*
1-2-R 206	CGAGAGGCTGAAGCTATTAGTG	Exon 5 of *TIC110*
2-2-F 201	TTACCCTGATCAACTGGAGCTT	Exon 1 of *TOC132*
2-2-R 202	ACGGACAGAAGAAGAGGTTGTAG	Exon 1 of *TOC132*
3-1-F 195	ACCTTAGAATCCAGAGTTGGTG	Intron 7 of *Hsp93-III*
3-1-R 196	GCTTGGTCGATAGCTCTTCTTA	Intron 9 of *Hsp93-III*
4-2-F 193	GTCGCATCGGTTGATTCTTACT	Intron 1 of *TIC20-IV*
4-2-R 194	GTGCACCATATGACCTGAAGAG	Exon 3 of *TIC20-IV*
5-1-F 187	GTTGTGACCTGAGTCTGAACTG	Intron 5 of *Hsp93-V*
5-1-R 188	CAGCTCGAGTCCTTGAGAATTTAG	Exon 8 of *Hsp93-V*

on a different chromosome, from promising members of the TIC/TOC chloroplast import system: (1) *Tic110*; (2) *Toc132*; (3) *Hsp93-III*; (4) *Tic20-IV*; and (5) *Hsp93-V*. Each of these loci was chosen based on location in the Arabidopsis genome and potential interactions with ACC2. The different chromosomal locations of these loci allowed us to test both the gene itself as a potential modifier, and the region around that gene for linkage to a potential modifier (Figure 26). We initially focused on members of the TIC/TOC system because transport of ACC2 into the chloroplast is required in order to compensate for loss of the heteromeric ACCase protein. As described in Chapter 2, TIC110 along with the two chaperone proteins, Hsp93-III and Hsp93-V, function in the translocation motor that guides proteins such as ACC2 into the chloroplast stroma (Kovacheva et al., 2005; Shi and Theg, 2013). cpHsc70-2, another chaperone of the translocation motor, was not tested due to its close proximity to Hsp93-V on chromosome 5. If TIC110 is a modifier, a change in the protein might have a downstream effect on the import of ACC2 into the chloroplast through its function as a recruiter of stromal chaperone proteins such as Hsp93-III and Hsp93-V (Kovacheva et al., 2005). Loss of one of these chaperone proteins, either through failed recruitment by TIC110 or a mutation in the protein itself, would likely affect the folding and stability of the precursor protein as it is moved into the stroma.

TOC132, along with its partner TOC120, functions similarly to move ACC2 and other housekeeping proteins across the outer membrane (Hirabayashi et al., 2011; Shi and Theg, 2013). TOC132 is a more likely modifier candidate than TOC120 due to the A-domain within TOC132, which functions in the initial recognition of the transit peptide sequences of chloroplast-localized housekeeping proteins (Inoue et al., 2010). A change in the A-domain of TOC132 would likely affect the recruitment of ACC2 to the TOC import system. Within the region surrounding *Toc132* on chromosome 2 is *Tic21*, whose protein product likely helps to assemble the 1-MDa import complex on the inner membrane (Teng et al., 2006; Shi and Theg, 2013). TIC20-IV is thought to

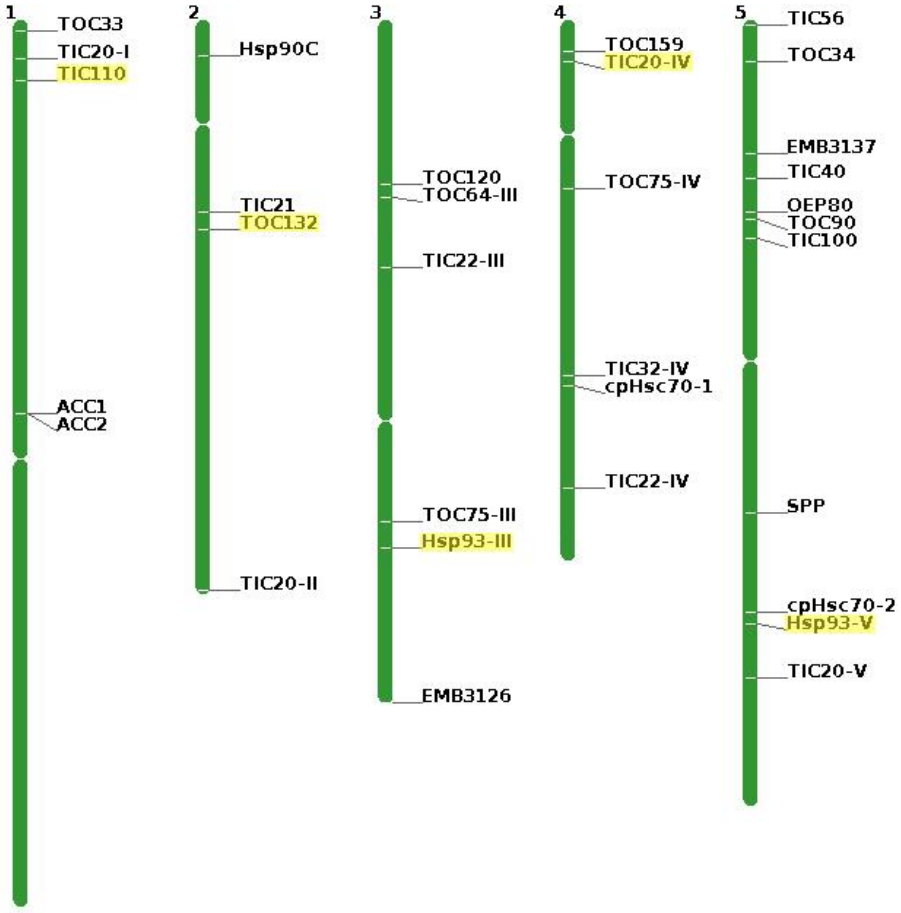


Figure 26. Chromosome Locations of Arabidopsis Genes Encoding Known Components of the Chloroplast Protein Import System. The highlight genes are the five loci genotyped in our candidate gene approach. *ACC1* and *ACC2* are located near the centromere on chromosome 1, *EMB3126* is located near the bottom of chromosome 3, and *EMB3137* is located near the top of chromosome 5. Adapted from Parker et al. (2014).

play a crucial role as the main channel protein through which housekeeping proteins are moved across the inner membrane (Hirabayashi et al., 2011; Kasmati et al., 2011; Kikuchi et al., 2013). The *tic20-iv* knockout mutant (SAIL_97_F10) grows normally on basal medium and soil, which means that there is redundancy in the TIC import system where loss of one channel protein is compensated by another. In this case, the redundant protein is likely TIC20-I, the main channel for import of photosynthetic proteins (Kikuchi et al., 2013), or one of the two TIC20 proteins whose functions are not known: TIC20-II and TIC20-V (Kasmati et al., 2011; Shi and Theg, 2013). Whereas there is redundancy in the transport of housekeeping proteins across the inner membrane, TIC20-IV seems to be the only channel protein for the movement of ACC2 into the stroma. This can be seen in the hypersensitive response of *tic20-iv* seedlings on spectinomycin, which is similar to that of a null mutant of *ACC2*. Under this model, a tolerant allele of *Tic20-IV* may increase the efficiency of ACC2 transport. All of these candidate loci were chosen prior to Kayla's work on the enhancer region.

Arabidopsis Progeny Lines Chosen for Analysis

Five descendent lines were chosen for this analysis from a cross between the tolerant Tsu-0 accession and *emb3126-1* mutant in the "Nossen" background (Table 27; Figures 27 and 28). All of these lines are homozygous for the Tsu-0 allele of *ACC2*, which means they contain a fully functional suppressor. They are also homozygous for the Tsu-0 allele of the enhancer on chromosome 5, as shown by the genotype of *EMB3137* and two surrounding genes: *Toc34* and *Oep80*. The descendent lines were divided into two groups based solely on the predicted genotypes of potential modifier loci. The first group (1B-3B-1A, 1B-3B-2E, and 20D-3A-2A) showed the highest level of rescue among all lines screened. These lines are therefore most likely to be homozygous Tsu-0 for any modifiers that increase embryo rescue. The second group

Table 27. Embryo Rescue in Progeny Plants Screened for Candidate Modifiers

Plant Name [a]	Mutant Seeds Screened	Embryo Lengths (µm)		Embryos Measured (%)			Embryos Stages (%)			
		Average [b]	*t*-value [c]	< 100 µm	> 100 µm	> 200 µm	Globular	Triangular	Linear	Cotyledon
1B-3B-1A	55	328 ± 11.2	-	100.0	94.5	54.5	0.0	0.0	29.1	70.9
1B-3B-1A-2C	33	319 ± 17.2	0.4	0.0	100.0	78.8	0.0	0.0	45.5	54.5
1B-3B-1A-2D	32	270 ± 13.5	3.0 **	0.0	100.0	75.0	0.0	0.0	53.1	46.9
1B-3B-2E	52	350 ± 13.2	-	100.0	100.0	61.5	0.0	0.0	36.5	63.5
1B-3B-2E-2B	46	273 ± 14.0	4.0 ***	0.0	100.0	67.4	0.0	0.0	65.2	34.8
1B-3B-2E-2E	28	286 ± 19.3	2.7 **	0.0	100.0	78.6	0.0	0.0	53.6	46.4
20D-3A-2A	38	274 ± 16.2	-	100.0	78.9	34.2	0.0	5.3	36.8	57.9
20D-3A-2A-2A	32	307 ± 13.0	-0.1	0.0	100.0	87.5	0.0	0.0	28.1	71.9
20D-3A-2A-2D	30	276 ± 20.4	-0.1	0.0	100.0	70.0	3.3	3.3	53.4	40.0
20D-3A-2A-2E	26	304 ± 20.0	-1.2	0.0	100.0	80.8	0.0	0.0	34.6	65.4
S2-10D-2B	51	106 ± 4.2	-	37.3	56.9	0.0	41.2	52.9	5.9	0.0
S2-10D-2B-2B	32	96 ± 2.1	2.1 *	43.8	25.0	0.0	96.9	3.1	0.0	0.0
S2-10D-2B-2D	37	112 ± 4.7	-1.0	21.6	54.1	2.7	64.9	29.7	5.4	0.0
S2-10D-2B-2E	29	118 ± 4.4	-2.0 *	13.8	65.5	0.0	48.3	44.8	6.9	0.0
S2-3B-2D	66	95 ± 5.1	-	54.5	19.7	1.5	83.3	9.1	6.1	1.5
S2-3B-2D-1A	30	121 ± 4.0	-4.0 ***	6.7	73.3	0.0	36.7	60.0	3.3	0.0
S2-3B-2D-1B	29	126 ± 4.9	-4.3 ***	10.3	72.4	0.0	27.6	55.2	13.8	3.4
S2-3B-2D-1E	21	130 ± 5.9	-4.4 ***	4.8	76.2	0.0	23.8	57.2	19.0	0.0

[a] Gray font, parental lines chosen for this analysis. Black font, progeny plants screened from each parental line.
[b] Mean Length ± Standard Error.
[c] This column gives the T-test results of each progeny line compared to the parent line it was harvested from. Asteriks denote the significance level: * $p < 0.05$; ** $p < 0.01$; ***$p < 0.001$.

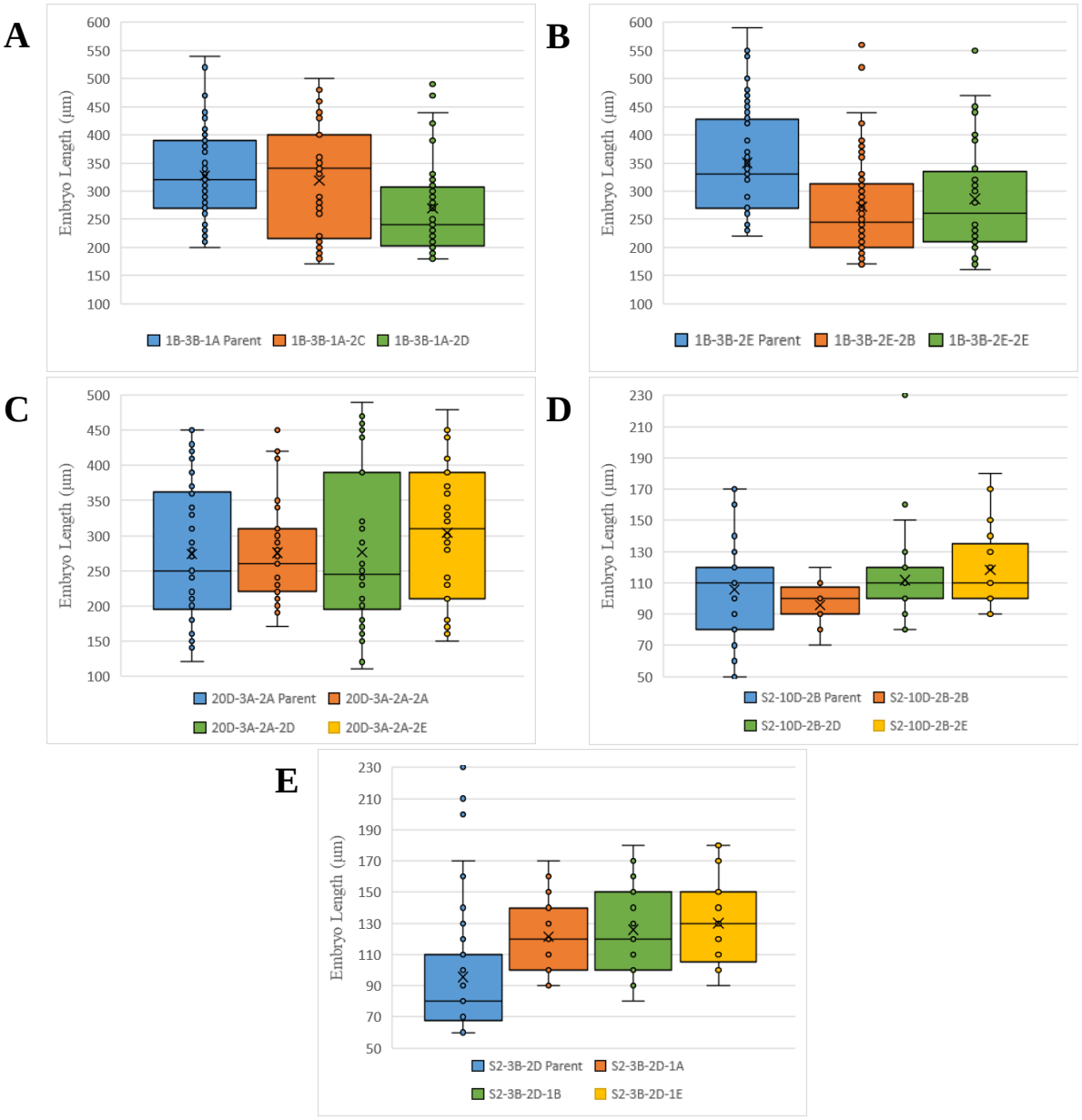


Figure 27. Boxplot Comparison of Mutant Embryo Length in Progeny Plants Screened for Candidate Modifiers and Their Parental Lines. Boxplots representing the median, 25th and 75th percentiles (interquartile range) of mutant embryo lengths. Whiskers extend to the minimum and maximum lengths (excluding outliers). Mean is denoted by the X. A, progeny of 1B-3B-1A; B. progeny of 1B-3B-2E; C, progeny of 20D-3A-2A; D, progeny of S2-10D-2B; and E, progeny of S2-3B-2D.

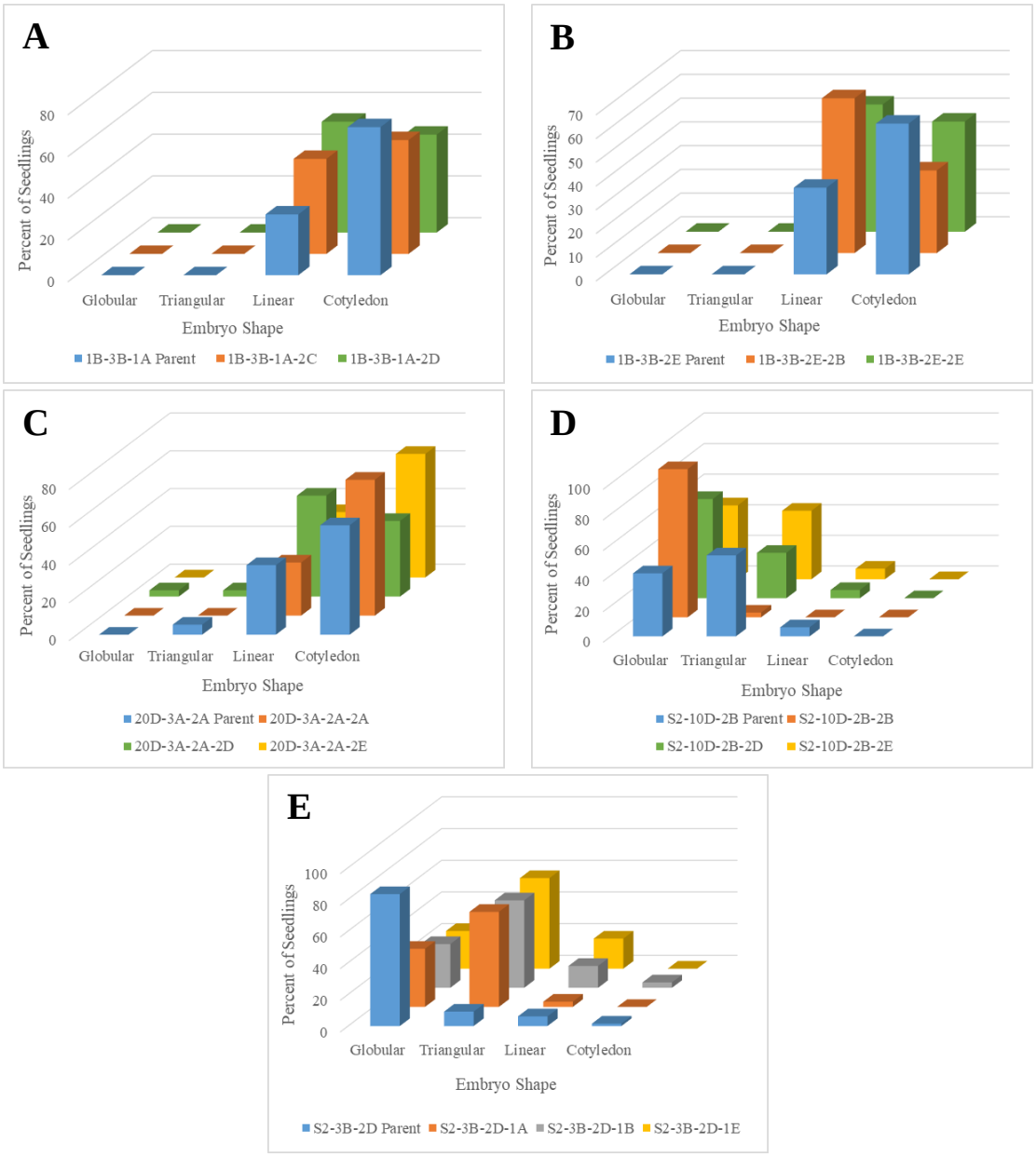


Figure 28. Comparison of Mutant Embryo Growth Stages in Progeny Plants Screened for Candidate Modifiers and Their Parental Lines. Percentage of embryos from each cross assigned to four phenotypic categories based on shape of the embryo: Globular, Triangluar, Linear, and Cotyledon. A, progeny of 1B-3B-1A; B. progeny of 1B-3B-2E; C, progeny of 20D-3A-2A; D, progeny of S2-10D-2B; and E, progeny of S2-3B-2D.

(S2-10D-2B and S2-3B-2D) exhibited the lowest level of embryo rescue while still having the Tsu-0 alleles of *ACC2* and the enhancer. This means that these two lines most likely have the "Nossen" allele of any modifiers affecting the amount of embryo rescue.

In order to keep track of the descendent lines used in this study, the name of each line consists of the plant identification number that dry seed was harvested from in each generation. For example, descendent line 1A-1B-1C derived from plant 1A in the F2 generation, 1B in the F3 generation, and 1C in the F4 generation. After plating progeny seed from a single F1 plant, F2 plants were screened for the amount of embryo rescue and genotyped for *ACC2* and the enhancer. From the F2 generation, plants 1B and 20D exhibited the most rescue, and were genotyped to have a Tsu-0 allele for *ACC2* and the enhancer. Even more advanced rescue was seen in the F3 generation, where plants 1B-3B and 20D-3A showed the most rescue. These plants still maintained the Tsu-0 genotype of *ACC2* and the enhancer. In the F4 generation, plant 20D-3A-2A exhibited the same amount of rescue as its parental F3, whereas plants 1B-3B-2E and 1B-3B-1A showed further increase in the average embryo rescue. All three of these F4 plants should be homozygous Tsu-0 for any potential modifiers tested in this analysis. During a second round of screening of F2 plants (labeled as S2), plants S2-10D and S2-3B showed limited embryo rescue, have a Tsu-0 allele of *ACC2*, and are heterozygous at the enhancer loci. In the F3 generation, plants S2-10D-2B and S2-3B-2D continued to show low amount of embryo rescue, but they were genotyped as Tsu-0 for both *ACC2* and the enhancer locus. Both of these F3 plants should be homozygous "Nossen" for any potential modifiers. Genomic DNA was extracted from progeny seedlings of all five descendent lines. Sibling seedlings grown at the same time, were transplanted to soil, grown to maturity, and were screened to confirm the amount of embryo rescue for each descendent line.

RESULTS

Table 27 shows the embryo phenotypes of progeny plants compared to the parental lines chosen for this analysis. Over half of the progeny screened were significantly different from their parental lines. No significance was found in the progeny plants of 20D-3A-2A (Figures 24 and 25, part C). This is possibly due to the wider spread in the mutant embryo lengths measured. In 1B-3B-1A and S2-10D-2B, there was a mixture of plants that were significantly different from their parent line and plants that were not. These slight differences in the progeny plants can be seen in in Figures 27 and 28 parts A and D. The progeny plants of 1B-3B-2E and S2-3B-2D all showed a significant difference from their parental lines. In 1B-3B-2E there is a decrease in the size of the embryos measured, and an increase in the number of embryos arresting at a lower (linear) stage of development rather than the cotyledon stage (Figures 27 and 28, part B). The opposite is happening with the progeny plants of S2-3B-2D: there is an increase in the embryo size and an increase in the number of embryos arresting at higher stages of development rather than the globular stage (Figures 27 and 28, part E). The differences found between the progeny plants and their parental lines could indicate heterozygosity of one or more modifier loci in the parental lines.

For each descendent line, three individual seedlings were PCR genotyped at each of the five loci, which gave a total of 15 progeny tested. Additional seedlings were available for analysis if a promising candidate was found. However, none of the results were consistent with the expected outcome if a modifier locus was linked to one of the candidates (Table 28). For complete linkage, we expected to see the group with the highest level of embryo rescue (1B-3B-1A, 1B-3B-2E, and 20D-2A-3A) homozygous for the Tsu-0 allele of the candidate modifier whereas the group with the lowest rescue (S2-10D-2B and S2-3B-2D) homozygous for the "Nossen" allele. In order to detect a locus linked to one of the candidates, we would expect that a low number, around one to three, of the 15 progeny seedlings tested would differ from the

predicted genotype. More than three or so differences would be unreliable for determining linkage using the small number of seedlings examined. If any loci had shown potential linkage, additional seedlings for each descendent line would have been tested to provide more accurate results. The expected results if none of the modifiers are linked to a candidate locus are harder to predict due to the locations of crossovers within each descendent line and the low possibility of heterozygosity at the locus.

Table 28 shows the genotype of each candidate locus for all 15 seedlings analyzed. The results of all loci for 1B-3B-1A and 1B-3B-2E were exactly the same. This is likely due to how closely related the two descendent lines are as they were both harvested from the same F3 plant. It is also likely that all five loci genotyped here were homozygous in the F3 plant (1B-3B). Heterozygous regions, as seen with four out of five loci (indicated in Table 28 by one to four "Het" seedlings), are expected for a minor percentage of the genome. This is due to the genome of each subsequent generation from a cross becoming more isogenic. For example, the use of seedlings from generations F8 and F9 would have led to results with significantly less heterozygosity than the F4 and F5 seedlings analyzed here. This effect can be seen in Table 28, where there is one example of a heterozygous region in the F5 progeny (1B-3B-1A, 1B-3B-2E, and 20D-2A-3A) and five examples in the F4 progeny (S2-10D-2B and S2-3B-2D). *Toc132* produced the most interesting results because the genotypes were the exact opposite of what we expected. All plants tested from the highest group were homozygous "Nossen" for *Toc132,* and all plants tested from the lowest group were homozygous Tsu-0. A possible explanation for this result is that TOC132 in "Nossen" is fully functional in recruiting and translocating ACC2 across the outer membrane of the chloroplast, whereas TOC132 in Tsu-0 has reduced function. However, it is most likely that this locus would show results similar to the other loci if additional descendent lines were tested.

Table 28. Genotypes of Candidate Modifier Loci for Each Descendent Line Tested

Parental Line	Embryo Phenotype	TIC110 [a] (Chromosome 1)	TOC132 (Chromosome 2)	Hsp93-III (Chromosome 3)	TIC20-IV (Chromosome 4)	Hsp93-V (Chromosome 5)
1B-3B-1A	Late	N, N, N	N, N, N	T, T, T	T, --, T	N, N, N
1B-3B-2E	Late	N, N, N	N, N, N	--, T, T	T, T, T	N, N, N
20D-3A-2A	Late	N, N, N	N, N, N	N, N, N	H, H, N	T, T, T
S2-10D-2B	Early	H, H, N	T, T, T	T, T, T	T, T, T	N, T, H
S2-3B-2D	Early	T, T, T	T, T, T	T, T, H	H, H, N	H, N, N

[a] Letters represent the genotype of three progeny seedlings tested for each parental line. N, homozygous "Nossen". T, homozygous Tsu-0. H, heterozygous. Red dashes, seedlings whose sequences could not be analyzed.

DISCUSSION

This chapter describes a candidate gene approach to test five members of the TIC/TOC chloroplast import system (*Tic110*, *Toc132*, *Hsp93-III*, *Tic20-IV*, and *Hsp93-V*) as potential modifiers that increase embryo rescue in the presence of a functional suppressor (*ACC2*) and tolerant enhancer. No correlation was found between the amount of embryo rescue and the genotype at these five loci. With the number of seedlings examined for each locus, only close linkage with the candidate gene can be detected. After that, the results are unreliable. This leaves large portions of the genome not examined in this analysis. Each chromosome in Arabidopsis is 76-122 cM in length (Meinke et al., 2009). This means that at most we are able to detect linkage across half of each chromosome. Another issue is the possibility that two or more modifiers interact, with each one partially contributing to the extent of embryo development seen in the group with the most rescue (1B-3B-1A, 1B-3B-2E, and 20D-2A-3A). In this case, the modifiers are acting similar to quantitative trait loci, which would make it difficult to identify an individual locus through a candidate gene approach. This also raises the question: Does the enhancer have a stronger effect on embryo rescue than a single genetic modifier? In other words, if we had used a different *emb* mutant in our initial crosses that was linked to one of the modifier loci, would we have considered that locus to be the enhancer? To answer these questions, we can compare the amount of embryo rescue between two groups of descendent lines: (1) those that are likely homozygous "Nossen" for the enhancer and homozygous Tsu-0 for the modifiers; and (2) those that are likely homozygous Tsu-0 for the enhancer and homozygous "Nossen" for the modifiers. If the enhancer alone has a stronger influence on embryo rescue than the modifiers, we would expect to see a higher level of rescue in Group 2 than in Group 1. This is exactly what we observed. The embryos rescued in Group 2 averaged around 104 μm (± 2.4, SE) in length whereas the embryos in Group 1 averaged around 84 μm (± 3.0, SE). This more substantial effect of the enhancer on embryo development is consistent with the requirement of a tolerant enhancer

for the modifiers to significantly extend embryo development to later stages. Because we still do not know which gene is the enhancer, it is difficult to build a model for the function of the modifiers. It might therefore be more beneficial to narrow down the enhancer locus before further attempts to identify the modifiers.

Prior to Kayla's work on the enhancer, we focused on members of the TIC/TOC system as potential modifiers because we suspected that many of these proteins interact with ACC2 to facilitate import into the chloroplast. Through our candidate gene approach, we were not able to show that a modifier locus is linked to any of our five candidates. The translocation motor, including TIC110, functions to transport all housekeeping and photosynthetic proteins into the chloroplast. Therefore, it is not surprising that *Tic110* is not a modifier because a change in the protein that affects import of ACC2 into the chloroplast would also likely affect other chloroplast-localized proteins. The chaperone proteins, Hsp93-III and Hsp93-V, are likely just as important in the translocation of housekeeping and photosynthetic proteins across the inner membrane, so a change that affects ACC2 import would also affect others. Similar to the translocation motor, the recognition by the A-domain of TOC132 is likely important for the import of many proteins into the chloroplast, not just ACC2. A change within this domain is likely to also affect the import of other housekeeping proteins. TIC20-IV seemed to be the most likely candidate due to the redundancy in the import of housekeeping proteins found through the normal growth of *tic20-iv* knockout mutants. The hypersensitivity of *tic20-iv* to spectinomycin also indicates that TIC20-IV is required for import of ACC2 into the chloroplast. However, the genotype results showed no linkage between *Tic20-IV* and the level of embryo rescue in the descendent lines. This does not rule out the possibility that a defect in *tic20-iv* is at least partially responsible for a decrease in tolerance to spectinomycin of another accession.

Several other candidate modifiers from the TIC/TOC system that are located in untested regions of the genome have not been evaluated (Figure 26): (1) *Toc120*, whose protein product

complexes with TOC132 to function as the main GTPase in transport of housekeeping genes across the outer membrane (Hirabayashi et al., 2011; Shi and Theg, 2013); (2) *Hsp90C*, which encodes another chaperone protein associated with the TIC110 translocation motor (Kovacheva et al., 2007; Inoue et al., 2013; Shi and Theg, 2013); (3) *cpHsc70-1*, a TIC110 translocation chaperone protein (Kovacheva et al., 2007; Inoue et al., 2013; Shi and Theg, 2013); and (4) TIC22-IV, which encodes a chaperone protein thought to guide precursor proteins between the TIC and TOC complexes within the intermembrane space (Kouranov et al., 1998; Shi and Theg, 2013).

In her work to identify candidates for the enhancer, Kayla manually curated 104 and 101 loci upstream and downstream, respectively, of *EMB3137*, which is closely linked to the enhancer (Cook and Meinke, 2017). She also did a quick scan of an additional 100 genes above and below this region for any obvious candidates. While trying to identify enhancer candidates, Kayla looked for proteins whose function would fall into one of our models for function of the enhancer, including potential interactions between the enhancer and ACC2 (Table 29). My work on identifying potential modifier loci focused on proteins that fell into Model 1c: the improvement of ACC2 import through chloroplast membrane. The two chaperone proteins I tested, Hsp93-III and Hsp93-V, might also function in stabilization, folding and dimerization of ACC2 once it has moved into the stroma (Model 1d).

Kayla identified seven candidate genes as potential enhancers, and ranked these genes based on how well they fit a model for the function of the enhancer. All seven candidates are described in Table 30. Of the three most promising candidates, two (*GUN5* and *NACA3*) function in protein complexes, and interact with other proteins that could be potential modifiers. *GUN5* encodes a subunit (CHLH) of the magnesium-protoporphyrin IX (Mg-ProtoIX) chelatase that functions in bacteriochlorophyll and chlorophyll biosynthesis and ABA signaling (Walker and Willows, 1997; Du et al., 2012). CHLH also has a second function in retrograde signaling

Table 29. Models for Enhancer Function in the Absence of Chloroplast Translation. Adapted from Cook and Meinke (2017).

1.		Enhances Function, Abundance or Localization of ACC2
	a.	Improves translational efficiency of *ACC2* mRNA
	b.	Improves targeting of ACC2 to plastid via chaperone molecule
	c.	Improves import of ACC2 through plastid membrane
	d.	Improves ACC2 folding and dimerization inside plastid
2.		Improves Fatty Acid Biosynthesis in Plastid
	a.	Increases efficiency of upstream/downstream reactions
	b.	Improves export of ACC2-synthesized fatty acids
3.		Compensates for Loss of Ycf1, Ycf2, ClpP1 Functions in Plastid
4.		Impacts Chloroplast-Nucleus Retrograde Signaling Pathways
5.		Improves Other Rate-Limiting Metabolic Pathways in Plastid

Table 30. Enhancer Candidates Identified in the Region Flanking *EMB3137*. Adapted from Cook and Meinke (2017).

Rank [a]	Locus Number	Gene Symbol	Edited Function [b]	Edited Function Details [b]
A	At5g13390	*NEF1*	Plastid Integral Membrane Protein	Required for pollen exine formation; Proposed roles in plastid membrane integrity and fatty acid export
A	At5g13630	*ABAR; CCH; CHLH; GUN5*	Magnesium Chelatase	Plastid to nucleus retrograde signal transduction
A	At5g13850	*NACA3*	Nascent Polypeptide Associated Complex Subunit Alpha-Like Protein 3	Potential role in translocation of nascent polypeptides into chloroplasts
A/B	At5g13410		Plastid-Localized FKBP-Like Protein; Immunophilin	Potential role in protein folding
A/B	At5g13640	*PDAT1*	Phospholipid: Diacylglycerol Acyltransferase	TAG biosynthesis; Fatty acid and membrane lipid homeostasis
A/B	At5g15450	*CLPB3;*	Plastid-Localized ClpB Homologue; Chaperone	Remodeling of protein aggregates
B	At5g12860	*DIT1; OMT1*	Plastid Dicarboxylate Transporter	Integration of carbon, nitrogen metabolism

[a] System used to subjectively rank each enhancer candidate locus. A, most likely; A/B, promising; B, possible.

[b] Based on information from TAIR (http://www.arabidopsis.org/) and relevant publications.

between the chloroplast and nuclear genomes (Mochizuki et al., 2001; Du et al., 2012). Kayla's model for *GUN5* as the enhancer is based on its secondary function in retrograde signaling where a tolerant (Tsu-0) version of the enhancer would limit the passage of a signal from the chloroplast genome when it is inhibited to the nuclear genome, which would allow the nuclear genes normally affected by the signal to continue to be expressed. If *GUN5* is the enhancer locus, potential modifiers in this system include *GUN4* (At3g59400), *CHLI1* (At4g18480), and *CHLD* (At1g08520), which all encode subunits of the Mg-ProtoIX chelatase (Figure 29; Du et al., 2012). As modifiers, these loci would function alongside GUN5 in retrograde signaling. However, this does not explain the low levels of embryo rescue seen when tolerant alleles of the modifiers are present, but the enhancer is sensitive. As modifiers, these three proteins would require functional GUN5 to be present. The Arabidopsis Interactions Viewer, which shows protein-protein interactions, indicates that GUN5 also interacts with SYP23 (At4g17730), which is involved in vesicle-mediated transport, and CKA4 (At2g23070), a chloroplast-localized subunit of casein kinase 4 (http://bar.utoronto.ca/interactions/cgi-bin/arabidopsis_interactions_viewer.cgi). CKA4 is thought to be involved in the same retrograde signaling pathway as the Mg-ProtoIX chelatase complex containing GUN5 (Wang et al., 2014). In other words, as a modifier CKA4 would function similar to the other members of the complex. SYP23 as a modifier could possibly be involved in transport of signaling molecules within the retrograde signaling pathway involving Mg-ProtoIX Chelatase and CKA4 along with other molecules in the cell, which could explain the slight increase in embryo rescue when the enhancer (GUN5) is not present.

The other promising candidate for the enhancer locus (*NACA3*), studied most extensively in yeast, is thought to encode the alpha subunit of the Nascent Polypeptide Associated Complex (NAC; Ponce-Rojas et al., 2017). This complex functions as a chaperone for newly synthesized polypeptide chains including the translocation of these precursor proteins to both the mitochondria and the chloroplast (Yang et al., 2007). There is also evidence for independent

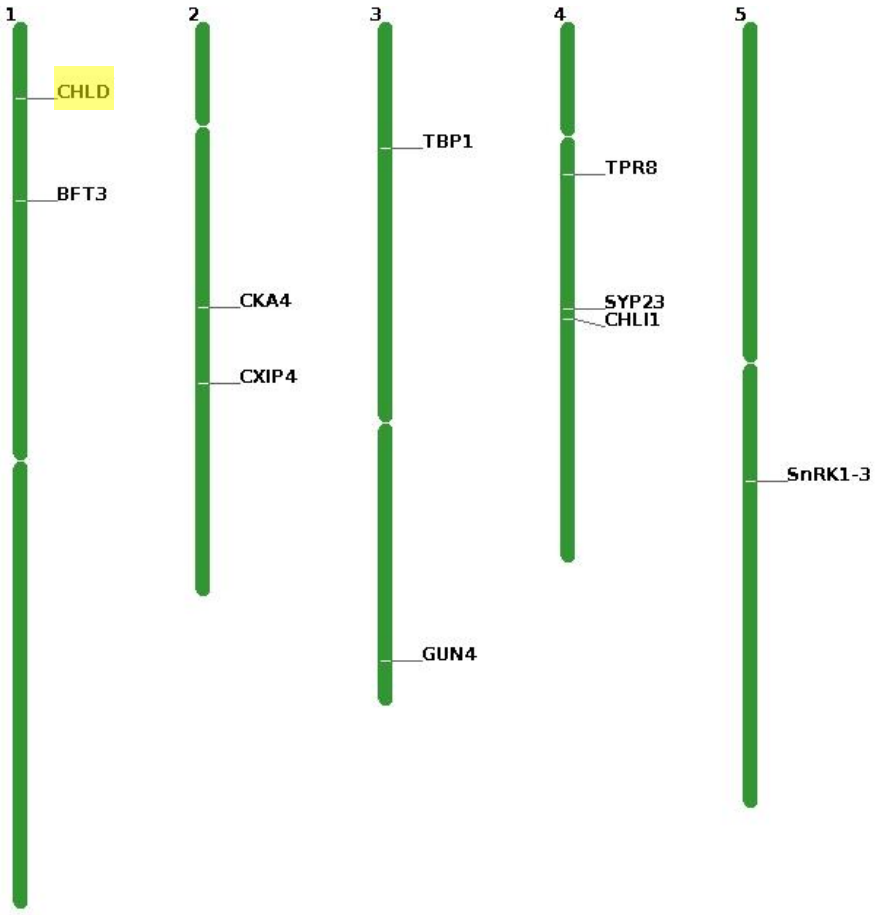


Figure 29. Chromosome Locations of Arabidopsis Genes Encoding Untested Modifier Candidates. If *GUN5* is the enhancer, potential modifiers are *CHLD*, *CKA4*, *GUN4*, *SYP23*, and *CHLI1*. If NACA3 is the enhancer, potential modifiers are *BTF3*, *CXIP4*, *TBP1*, *TPR8*, and *SnRK1-3*. *CHLD* is highlighted to show that it is unlikely to be a potential modifier, because it is closely linked to *Tic110*, which has been tested.

function of NACA3 as a transcription factor (Moreau et al., 1998; Yang et al., 2007). Kayla's model for *NACA3* as the enhancer is that the NAC functions in translocation of ACC2 to the chloroplast membrane. A tolerant version of NACA3 might increase the efficiency of ACC2 targeting to the membrane. According to this model, potential modifiers would include the beta subunit of NAC, *BTF3* (At1g17880; Figure 29), which would function alongside NACA3 in translocation of ACC2 precursor proteins to the chloroplast membrane. As for the slight increase in embryo rescue in the absence of the enhancer (NACA3), BTF3 has also been shown in humans and *C. elegans* to function as a transcription factor and a suppressor of apoptosis independently of NAC (Yang et al., 2007). Additional proteins that interact with NACA3 include CXIP4 (At2g28910), which regulates calcium transport, TBP1 (At3g13445), a transcription factor that binds to the TATA box promoter region, TPR8 (At4g08320), a tetratricopeptide repeat protein with unknown function, and SnRK1-3 (At5g39440), a phosphorylase (http://bar.utoronto.ca/interactions/cgi-bin/arabidopsis_interactions_viewer.cgi). These proteins likely interact with NACA3 in its role as an activator of C-Jun-dependent transcription, where it interacts with a TATA box binding protein, which is likely TBP1, and a phosphorylase, which is likely SnRK1-3 (Moreau et al., 1998). CXIP4 could possibly function in activation of NACA3 through a calcium signaling pathway. The slight embryo rescue seen when the modifiers are present and the enhancer is not could be explained through transcription regulation activities of these proteins that do not involve NACA3.

FUTURE DIRECTIONS

Much has been accomplished with this project towards understanding why plant species differ in their ability to tolerate a loss of chloroplast translation. Through studies at both the embryo and seedling stages, we found that functional alleles of *ACC2* can suppress the

preglobular phenotype of *emb* mutants defective in chloroplast translation and the sensitive phenotype of seedlings grown on spectinomycin. We also identified a locus on chromosome 5 that enhances the suppressor effect of *ACC2*, and uncovered evidence for additional modifiers that further rescue the embryo and seedling phenotypes. Additional studies using the natural variation found in Arabidopsis accessions looked at the effects of mutations on the structure and function of ACCase proteins. Through this work, we have identified a number of null mutations that eliminate ACC2 function, and some missense mutations whose effects range from partial to severe loss of ACC2 function. However, there are still areas within this project where additional work can be done to help further our understanding of this system.

One future area to address would be to search for additional natural variation in *ACC1* and *ACC2* sequences. In the time since we obtained the sequences from 855 accessions through the 1001 Genomes Project, the genome sequences of 280 more accessions have been published (The 1001 Genomes Consortium, 2016). Even though these newer sequences are not available through the Salk Genome Browser (http://signal.salk.edu/atg1001), there are new sequence viewers on the 1001 Genomes Project website that incorporate all 1,135 sequences. The 1001 Proteomes viewer (http://1001proteomes.masc-proteomics.org/) shows all non-synonymous single nucleotide polymorphisms (nsSNPs) at any locus of interest, whereas the Polymorph 1001 viewer (http://tools.1001genomes.org/polymorph/) shows all SNPs, insertions and deletions at any locus. The most important variation to identify for this project would be additional examples of the three missense mutations that significantly reduce or eliminate ACC2 function and are limited to a single accession: G135E in Sav-0, F1206L in Aitba-1, and E1689G in Ts-1. Additional sensitive accessions with one of these variants would provide further evidence of their deleterious effect on ACC2 protein function. Another variant of interest is A376V, found only in the Col-0 accession, where it likely reduces ACC2 function to some extent. We could also look for new missense mutations affecting residues that were not analyzed through our previous

studies (85% of the total residues), especially in highly conserved residues (>95%) found through our multi-kingdom alignment of 667 ACCase sequences.

In addition to the expanded natural variation that can be analyzed, artificial variation can be introduced using recent advances in gene editing technologies to produce missense mutations that alter residues of interest. Focusing on the mutations that most likely reduce or eliminate ACC2 function (G135E, A376V, F1206L, and E1689G along with I404K and T1902K in the Knox-18 group of accessions), we could introduce these mutations individually into the *ACC2* sequence of a tolerant accession like Tsu-0, and measure the effects on ACC2 function by looking for increased sensitivity on spectinomycin. More subtle changes in ACCase function could be measured by introducing each missense mutation into the *ACC1* sequence of Col-0, and comparing the embryo phenotypes with known *acc1* mutants (*emb22*, *pas3-1*, and *pas3-2*) whose terminal phenotype is determined by the strength of the mutation (Parker et al., 2016). Additional regions of interest for gene editing include the 17 other ACC2 variants found in natural accessions that seem to slightly reduce the function of ACC2. We could also use gene editing to analyze the effects of modifying the most conserved residues (>99%) from our multi-kingdom alignment where there is no natural variation to be evaluated. Gene editing technologies could also be used to analyze the effects of these missense mutations on plants that have non-functional copies of the enhancer and modifiers. In order to evaluate this further, we could utilize the many descendent lines from our cross between Tsu-0 and *emb3126-1*, and compare the change in spectinomycin sensitivity when the mutation is introduced. Descendent lines 1B-3B-1A and 20D-3A-2A, which contain Tsu-0 alleles of *ACC2*, the enhancer, and modifiers, could be compared to lines 3B-1A-1A and S2-3B-6B, which contain a Tsu-0 allele of *ACC2*, a "Nossen" allele of the enhancer, and likely "Nossen" alleles for the modifiers. Other comparisons could be made singling out just the enhancer, and just the modifiers.

Through our crosses between sensitive accessions and knockout mutants of *acc2* and

tic20-iv, we identified four accessions (La-0, Etna-2, Grivo-1, and Qar-8a) whose sensitivity seems to be caused by an unknown locus. This likely answers our question of whether there are other genes that give rise to a sensitive phenotype when disrupted. La-0 and Etna-2 are the most logical accessions to study first. In La-0, there is no obvious mutation in the *ACC2* or *Tic20-IV* sequences that would likely lead to sensitivity, as shown by Yixing Wang when she sequenced the full-length cDNA of *ACC2*. Results of La-0 crosses with knockout mutants clearly showed that neither gene led to the sensitivity of the accession. Results from Etna-2 crossed with the *acc2* knockout mutant were less definitive, but Etna-2 is the most sensitive of the four accessions.

One potential method to identify the locus responsible for sensitivity in La-0 and Etna-2 is to cross these accessions with a descendent line from our Tsu-0 x *emb3126-1* population that contains a functional (Tsu-0) allele of *ACC2* along with non-functional ("Nossen") alleles of the enhancer and modifiers, and use a mapping approach with markers spread throughout the genome. Using descendent lines that have sensitive alleles of the enhancer and modifiers (3B-1A-1A and S2-3B-6B) eliminates the effects those loci might have on the sensitivity of La-0 and Etna-2. This mapping approach would consist of PCR genotyping 50 tolerant F2 seedlings from these crosses for 15-20 markers equally spread across the genome. The focus here is on tolerant seedlings because Yixing Wang previously showed that it is difficult to isolate enough DNA from sensitive seedlings for multiple rounds of PCR genotyping, and tolerant F2 seedlings would not be homozygous La-0 or Etna-2 for the locus causing sensitivity. Individual or pooled sensitive F2 seedlings, which would be homozygous La-0 or Etna-2, could be used to confirm any candidate regions found.

A second approach for identifying the locus causing sensitivity would be to use the next generation mapping method developed by Austin et al. (2011). The same crosses between sensitive accessions (La-0 and Etna-2) and descendent lines lacking a functional enhancer and modifiers (3B-1A-1A and S2-3B-6B) could be used. For this approach, extracted DNA from 80-

100 sensitive F2 seedlings would be pooled and subjected to next-generation sequencing. The SNP frequencies across each chromosome would then be analyzed to find a non-recombinant region with low frequencies of polymorphism. Within this region should lie the locus (and mutation) responsible for the sensitivity of La-0 and Etna-2. Austin et al. (2011) have developed a method using a discordant chastity (ChD) statistic to further narrow the location of the responsible mutation by differentiating between causative mutations and SNPs that are likely due to natural variation. After next-generation sequencing of pooled DNA from F2 sensitive seedlings, the data can be uploaded and analyzed through the Next-Generation EMS Mutation Mapping website (http://bar.utoronto.ca/NGM/).

In addition to further understanding the function of ACCases, and looking at other genes that cause sensitivity in natural accessions, there is still more work to be done to identify the enhancer locus and potential modifiers. Kayla's work on identifying enhancer candidates could be expanded by PCR genotyping existing recombinant lines between *EMB3137, Toc34* (upstream) and *Oep80* (downstream) with additional markers within both of these regions. This would allow us to localize the enhancer to either upstream or downstream of *EMB3137*. Once this region is better defined, additional manual curation could be used to identify other potential candidates not found through Kayla's study. The "Nossen" genomic sequence for this smaller region of interest could also be obtained from Dr. Masatomo Kobayashi's lab at the RIKEN Plant Science Center, and then be used in sequence comparisons of candidate genes between Tsu-0, which has a functional enhancer, and "Nossen", which has a non-functional enhancer, to look for potential deleterious mutations. As more information is gained about the enhancer, additional candidate genes will arise as potential modifiers.

Because the original candidate gene approach to identify potential modifiers showed no linkage to the five genes chosen, either the candidate gene approach could be expanded with different genes located elsewhere in the genome, as discussed previously, or whole genome

sequences could be compared between descendent lines of Tsu-0 x *emb3126-1* that either have Tsu-0 (functional) or "Nossen" (non-functional) alleles for any potential modifiers. Similar to the previous candidate gene approach, we would be looking for regions of the genome where the lines with the least amount of embryo rescue (S2-10D-2B and S2-3B-2D) are homozygous "Nossen" whereas the lines with the most rescue (1B-3B-1A and 20D-3A-2A) are homozygous or heterozygous Tsu-0. Comparisons between other descendent lines that differ in only the functionality of the modifiers would help identify regions of interest.

Overall, the project described throughout this dissertation utilized natural variation in Arabidopsis accessions to study the effects of mutations, especially deleterious mutations, on a protein (ACCase) that is essential for fatty acid biosynthesis in eukaryotes. We also developed an understanding of some of the mechanisms behind the diverse phenotypic responses plant species have when translation of the chloroplast genome is blocked. Furthermore, our identification of accessions hypersensitive to spectinomycin has led to a more efficient method for plastid transformation in Arabidopsis (Yu et al., 2017).

REFERENCES

Abu-Elheiga L, Matzuk MM, Kordari P, Oh W, Shaikenov T, Gu Z, Wakil SJ (2005) Mutant mice lacking acetyl-CoA carboxylase 1 are embryonically lethal. Proc Natl Acad Sci USA **102**: 12011-12016

AGI (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature **408**: 796-815

Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, et al (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. Science **301**: 653-657

Alonso-Blanco C, Peeters AJ, Koorneef M, Lister C, Dean C, van den Bosch N, Pot J, Kuiper MT (1998) Development of an AFLP based linkage map of Ler, Col and Cvi *Arabidopsis thaliana* ecotypes and construction of a Ler/Cvi recombinant inbred line population. Plant J **14**: 259-271

Alonso-Blanco C, Aarts MG, Bentsink L, Keurentjes JJB, Reymond M, Vreugdenhil D, Koorneef M (2009) What has natural variation taught us about plant development, physiology, and adaptation? Plant Cell **21**: 1877-1896

Amid A, Lytovchenko A, Fernie AR, Warren G, Thorlby GJ (2012) The sensitive to freezing3 mutation of Arabidopsis thaliana is a cold-sensitive allele of homomeric acetyl-CoA carboxylase that results in cold-induced cuticle deficiencies. J Exp Bot **63**: 5289-5299

Asakura Y, Barkan A (2006) Arabidopsis orthologs of maize chloroplast splicing factors promote splicing of orthologous and species-specific group II introns. Plant Physiol **142**: 1656-1663

Austin RS, Vidaurre D, Stamatiou G, Breit R, Provart NJ, Bonetta D, Zhang J, Fung P, Gong Y, Wang PW, et al. (2011) Next-generation mapping of Arabidopsis genes. Plant J **67**: 715-725

Babiychuk E, Vandepoele K, Wissing J, Garcia-Diaz M, De Rycke R, Akbari H, Joubès J, Beeckman T, Jänsch L, Frentzen M, et al (2011) Plastid gene expression and plant development require a plastidic protein of the mitochondrial transcription termination factor family. Proc Nat Acad Sci USA **108**: 6674-6679

Baldwin A, Wardle A, Patel R, Dudley P, Park SK, Twell D, Inoue K, Jarvis P (2005) A molecular-genetic study of the Arabidopsis Toc75 gene family. Plant Physiol **138**: 715-733

Bates PD, Johnson SR, Cao X, Li J, Nam JW, Jaworski JG, Ohlrogge JB, Browse J (2014) Fatty acid synthesis is inhibited by inefficient utilization of unusual fatty acids for glycerolipid assembly. Proc Natl Acad Sci USA **111**: 1204-1209

Baud S, Guyon V, Kronenberger J, Wuillème S, Miquel M, Caboche M, Lepiniec L, Rochat C (2003) Multifunctional acetyl-CoA carboxylase 1 is essential for very long chain fatty acid elongation and embryo development in *Arabidopsis*. Plant J **33**: 75-86

Baud S, Bellec Y, Miquel M, Bellini C, Caboche M, Lepiniec L, Faure JD, Rochat C (2004) *gurke* and *pasticcino3* mutants affected in embryo development are impaired in acetyl-CoA carboxylase. EMBO Rep **5**: 515-520

Bechtold N, Ellis J, Pelletier, G (1993) In-planta Agrobacterium-mediated gene-transfer by infiltration of adult *Arabidopsis thaliana* plants. C. R. Acad. Sci. III **316**: 1194-1199

Bell CJ, Ecker JR (1994) Assignment of 30 microsatellite loci to the linkage map of Arabidopsis. Genomics **19**: 137-144

Berg M, Rogers R, Muralla R, Meinke D (2005) Requirement of aminoacyl-tRNA synthetases for gametogenesis and embryo development in Arabidopsis. Plant J **44**: 866–878

Bilder P, Lightle S, Bainbridge G, Ohren J, Finzel B, Sun F, Holley S, Al-Kassim L, Spessard C, Melnick M, et al. (2006) The structure of the carboxyltransferase component of acetyl-CoA carboxylase reveals a zinc-binding motif unique to the bacterial enzyme. Biochemistry **45**: 1712-1722

Bowman JL, Smyth DR, Meyerowitz EM (1991) Genetic interactions among floral homeotic genes of Arabidopsis. Development **112**: 1-20

Brodersen DE, Clemons WM, Carter AP, Morgan-Warren RJ, Wimberly BT, Ramakrishnan V (2000) The structural basis for the action of the antibiotics tetracycline, pactamycin, and hygromycin B on the 30S ribosomal subunit. Cell **103**: 1143-1154

Bryant N, Lloyd J, Sweeney C, Myouga F, Meinke D (2011) Identification of nuclear genes encoding chloroplast-localized proteins required for embryo development in Arabidopsis. Plant Physiol **155**: 1678-1689

Cai Z, Guisinger M, Kim HG, Ruck E, Blazier JC, McMurtry V, Kuehl JV, Boore J, Jansen RK (2008) Extensive reorganization of the plastid genome of *Trifolium subterraneum* (Fabaceae) is associated with numerous repeated sequences and novel DNA insertions. J Mol Evol **67:** 696-704.

Campbell JW, Cronan JE (2001) Bacterial fatty acid biosynthesis: targets for antibacterial drug discovery. Ann Rev Microbiol **55**: 205-332.

Carter AP, Clemons WM, Brodersen DE, Morgan-Warren RJ, Wimberly BT, Ramakrishnan V (2000) Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. Nature **407**: 340-348

Chalupska D, Lee HY, Faris JD, Evrard A, Chalhoub B, Haselkorn R, Gornicki P (2008) *Acc* homoeoloci and the evolution of wheat genomes. Proc Natl Acad Sci USA **105**: 9691-9696

Chang C, Bowman JL, Dejohn AW, Lander ES, Meyerowitz EM (1988) Restriction Fragment Length Polymorphism linkage map for *Arabidopsis thaliana*. Proc Natl Acad Sci USA **85**: 6856-6860

Cheng CY, Krishnakumar V, Chan AP, Thibaud-Nissen F, Schobel S, Town CD (2017) Araport11: a complete reannotation of the Arabidopsis thaliana reference genome. Plant J **89**: 789-804

Cheng F, Liu S, Wu J, Fang L, Sun S, Lie B, Li P, Hua W, Wang X (2011) BRAD, the genetics and genomics database for Brassica plants. BMC Plant Biol **11**: 136

Chory J, Ecker JR, Briggs S, Caboche M, Coruzzi GM, Cook D, Dangl J, Grant S, Guerinot ML, Henikoff S, et al (2000) National Science Foundation-sponsored workshop report: "The 2010 Project" – Functional genomics and the virtual plant. A blueprint for understanding how plants are built and how to improve them. Plant Physiol **123**: 423-426

Chou ML, Chu CC, Chen LJ, Akita M, Li HM (2006) Stimulation of transit-peptide release and ATP hydrolysis by a cochaperone during protein import into chloroplasts. J Cell Biol **175**: 893-900

Chumley TW, Palmer JD, Mower JP, Fourcade HM, Calie PJ, Boore JL, Jansen RK (2006) The complete chloroplast genome sequence of *Pelargonium x hortorum*: Organization and evolution of the largest and most highly rearranged chloroplast genome of land plants. Mol

Biol Evol **23**: 2175-2190

Clark R M, Schweikert G, Toomajian C, Ossowski S, Zeller G, Shinn P, Warthmann N, Hu TT, Fu G, Hinds DA, et al. (2007) Common sequence polymorphisms shaping genetic diversity in *Arabidopsis thaliana*. Science **317**: 338-342

Clough SJ, Bent AF (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. Plant J **16**: 735-743

Cook K, Meinke D (2017) A Curation-Based Search for the Enhancer Locus that Increases Tolerance of a Loss of Chloroplast Translation in *Arabidopsis thaliana*. Senior Honors Thesis, Oklahoma State University.

Constan D, Patel R, Keegstra K, Jarvis P (2004) An outer envelope membrane component of the plastid protein import apparatus plays an essential role in Arabidopsis. Plant J **38**: 93-106

Curci PL, Sonnante G (2016) The complete chloroplast genome of *Cynara humilis*. Mitochondrial DNA Part A **27**: 2345-2346

Daniell H, Lin CS, Yu M, Chang WJ (2016) Chloroplast genomes: diversity, evolution, and applications in genetic engineering. Genome Biol **17**: 134

dePamphilis CW, Palmer JD (1990) Loss of photosynthetic and chlororespiratory genes from the plastid genome of a parasitic flowering plant. Nature **348**: 337-339

Dinos G, Wilson DN, Teraoka Y, Szaflarski W, Fucini P, Kalpaxis D, Nierhaus KH (2004) Dissecting the ribosomal inhibition mechanisms of edeine and pactamycin: the universally conserved residues G693 and C795 regulate P-Site RNA binding. Molecular Cell **13**: 113-124

Douthwaite S (1992) Interaction of the antibiotics clindamycin and lincomycin with Escherichia coli 23S ribosomal RNA. Nucleic Acids Res **20**: 4717-4720

Drescher A, Ruf S, Calsa T Jr, Carrer H, Bock R (2000) The two largest chloroplast genome-encoded open reading frames of higher plants are essential genes. Plant J **22**: 97-104

Du S, Zhang X, Lu Z, Xin Q, Wu Z, Jiang T, Lu Y, Wang X, Zhang D (2012) Roles of the different components of magnesium chelatase in abscisic acid signal transduction. Plant Mol Biol **80**: 519-537

Dudas B, Jenes B, Kiss GB, Maliga P (2012) Spectinomycin resistance mutations in the rrn16 gene are new plastid markers in *Medicago sativa*. Theor Appl Genet **125**: 1517-1523

Duy D, Wanner G, Meda AR, Wirén NV, Soll J, Philippar K (2007) PIC1, an ancient permease in Arabidopsis chloroplasts, mediates iron transport. Plant Cell **19**: 986-1006

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res **32**: 1792-1797

Feldmann KA, Marks MD (1987) Agrobacterium-mediated transformation of germinating seeds of Arabidopsis thaliana: a non-tissue culture approach. Mol Gen Genet **208**: 1-9

Flores-Pérez Ú, Jarvis P (2013) Molecular chaperone involvement in chloroplast protein import. Biochim Biophys Acta **1833**: 332-340

Fu PC, Zhang YZ, Geng HM, Chen SL (2016) The complete chloroplast genome sequence of *Gentiana lawrencei* var. *farreri* (Gentianaceae) and comparative analysis with its congeneric species. PeerJ **4**: e2540

Goddard AD, Stevens JM, Rondelet A, Nomerotskaia E, Allen JW, Ferguson SJ (2010) Comparing the substrate specificities of cytochrome c biogenesis Systems I and II: bioenergetics. FEBS J **277**: 726-737

Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, et al (2012) Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res **40**: D1178-D1186

Gu K, Chiam H, Tian D, Yin Z (2011) Molecular cloning and expression of heteromeric ACCase subunit genes from *Jatropha curcas*. Plant Sci **180**: 642-649

Guisinger MM, Chumley TW, Kuehl JV, Boore JL, Jansen RK (2010) Implications of the plastid genome sequence of *Typha* (Typhaceae, Poales) for understanding genome evolution in Poaceae. J Mol Evol **70**: 149-166

Hager M, Biehler K, Illerhaus J, Ruf S, Bock R (1999) Targeted inactivation of the smallest plastid genome-encoded open reading frame reveals a novel and essential subunit of the cytochrome b(6)f complex. EMBO J **18**: 5834-5842

Haberle RC, Fourcade HM, Boore JL, Jansen RK (2008) Extensive rearrangements in the chloroplast genome of *Trachelium caeruleum* are associated with repeats and tRNA genes. J Mol Evol **66**: 350-361

Hiltbrunner A, Grünig K, Alvarez-Huerta M, Infanger S, Bauer J, Kessler F (2004) AtToc90, a new GTP-binding component of the *Arabidopsis* chloroplast protein import machinery. Plant Mol Biol **54**: 427-440

Hirabayashi Y, Kikuchi S, Oishi M, Nakai M (2011) In vivo studies on the roles of two closely related Arabidopsis Tic20 proteins, AtTic20-I and AtTic20-IV. Plant Cell Physiol **52**: 469-478

Hoja U, Marthol S, Hofmann J, Stegner S, Schulz R, Meier S, Greiner E, Schweizer E (2004) HFA1 encoding an organelle-specific acetyl-CoA carboxylase controls mitochondrial fatty acid synthesis in *Saccharomyces cerevisiae*. J Biol Chem **279**: 21779-21786

Hu TT, Pattyn P, Bakker EG, Cao J, Cheng JF, Clark RM, Fahlgren N, Fawcett JA, Grimwood J, Gundlach H, et al (2011) The *Arabidopsis lyrata* genome sequence and the basis of rapid genome size change. Nat Genet **43**: 476-481

Huang W, Ling Q, Bédard J, Lilley K, Jarvis P (2011) In vivo analyses of the roles of essential Omp85-related proteins in the chloroplast outer envelope membrane. Plant Physiol **157**: 147-159

Huang Y, Cho ST, Haryono M, Kuo CH (2017) Complete chloroplast genome sequence of common bermudagrass (*Cynodon dactylon* (L.) Pers.) and comparative analysis within the family Poaceae. PLoS One **12**: e0179055

Huerlimann R, Zenger KR, Jerry DR, Heimann K (2015) Phylogenetic analysis of nucleus-encoded acetyl-CoA carboxylases targeted at the cytosol and plastid of algae. PLoS One **10:** e0131099

Infanger S, Bischof S, Hiltbrunner A, Agne B, Baginsky S, Kessler F (2011) The chloroplast import receptor Toc90 partially restores the accumulation of Toc159 client proteins in the Arabidopsis thaliana ppi2 mutant. Mol Plant **4**: 252-263

Inoue H, Rounds C, Schnell DJ (2010) The molecular basis for distinct pathways for protein import into *Arabidopsis* chloroplasts. Plant Cell **22**: 1947-1960

Inoue H, Li M, Schnell DJ (2013) An essential role for chloroplast heat shock protein 90 (Hsp90C) in protein import into chloroplasts. Proc Natl Acad Sci USA **110**: 3173-3178

Jansen RK, Raubeson LA, Boore JL, dePhamphilis CW, Chumley TW, Haberle RC, Wyman SK, Alverson AJ, Peery R, Herman SJ, et al (2005) Methods for obtaining and analyzing whole chloroplast genome sequences. Methods Enzymol **395**: 348-384

Jansen RK, Cai Z, Raubeson LA, Daniell H, dePamphilis CW, Leebens-Mack J, Müller KF, Guisinger-Bellian M, Haberle RC, Hansen AK, et al (2007) Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. Proc Natl Acad Sci USA **104**: 19369-19374

Jansen RK, Wojciechowski MF, Sanniyasi E, Lee SB, Daniell H (2008) Complete plastid genome sequence of the chickpea (*Cicer arietinum*) and the phylogenetic distribution of *rps12* and *clpP* intron losses among legumes (Leguminosae). Mol Phylogenet Evol **48**: 1204-1217

Jarvis P (2008) Targeting of nucleus-encoded proteins to chloroplasts in plants. New Phytol **179**: 257-285

Johanson U, Hughes D (1995) A new mutation in 16S rRNA of *Escherichia coli* conferring spectinomycin resistance. Nucleic Acids Res **23**: 464-466

Kagale S, Robinson SJ, Nixon J, Xiao R, Huebert T, Condie J, Kessler D, Clarke WE, Edger PP, Links MG, et al. (2014) Polyploid Evolution of the Brassicaceae during the Cenozoic Era. Plant Cell **26**: 2777-2791

Kajiwara T, Furutani M, Hibara K, Tasaka M (2004) The *GURKE* gene encoding an acetyl-CoA carboxylase is required for partitioning the embryo apex into three subregions in Arabidopsis. Plant Cell Physiol **45**: 1122-1128

Kasmati AR, Töpel M, Patel R, Murtaza G, Jarvis P (2011) Molecular and genetic analyses of Tic20 homologues in *Arabidopsis thaliana* chloroplasts. Plant J **66**: 877-889

Kaundun SS (2014) Resistance to acetyl-CoA carboxylase inhibiting herbicides. Pest Manag Sci **70**: 1405-1417

Kessler F, Schnell D (2009) Chloroplast biogenesis: diversity and regulation of the protein import apparatus. Curr Opin Cell Biol **21**: 494-500

Kikuchi S, Oishi M, Hirabayashi Y, Lee DW, Hwang I, Nakai M (2009) A 1-megadalton translocation complex containing Tic20 and Tic21 mediates chloroplast protein import at the inner envelope membrane. Plant Cell **21**: 1781-1797

Kikuchi S, Bédard J, Hirano M, Hirabayashi Y, Oishi M, Imai M, Takase M, Ide T, Nakai M (2013) Uncovering the protein translocon at the chloroplast inner envelope membrane. Science **339**: 571-574

Kode V, Mudd EA, Iamtham S, Day A (2005) The tobacco plastid *accD* gene is essential and is required for leaf development. Plant J **44**: 237-244

Konieczny A, Ausubel FM (1993) A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers. Plant J **4**: 403-410

Koornneef M, Meinke D (2010) The development of Arabidopsis as a model plant. Plant J **61**: 909-921

Koornneef M, van Eden J, Hanhart CJ, Stam P, Braaksma FJ, Feenstra WJ (1983) Linkage map of *Arabidopsis thaliana*. J Hered **74**: 265-272

Kouranov A, Chen X, Fuks B, Schnell DJ (1998) Tic20 and Tic22 are new components of the protein import apparatus at the chloroplast inner envelope membrane. J Cell Biol **143**: 991-1002

Kovacheva S, Bédard J, Patel R, Dudley P, Twell D, Ríos G, Koncz C, Jarvis P (2005) *In vivo* studies on the roles of Tic110, Tic40 and Hsp93 during chloroplast protein import. Plant J **41**: 412-429

Kovacheva S, Bédard J, Wardle A, Patel R, Jarvis P (2007) Further *in vivo* studies on the role of the molecular chaperone, Hsp93, in plastid protein import. Plant J **50**: 364-379

Krech K, Ruf S, Masduki FF, Thiele W, Bednarczyk D, Albus CA, Tiller N, Hasse C, Schöttler MA, Bock R (2012) The plastid genome-encoded Ycf4 protein functions as a nonessential assembly factor for photosystem I in higher plants. Plant Physiol **159**: 579-591

Kubis S, Baldwin A, Patel R, Razzaq A, Dupree P, Lilley K, Kurth J, Leister D, Jarvis P (2003) The Arabidopsis *ppi1* mutant is specifically defective in the expression, chloroplast import, and accumulation of photosynthetic proteins. Plant Cell **15**: 1859-1871

Kubis S, Patel R, Combe J, Bédard J, Kovacheva S, Lilley K, Biehl A, Leister D, Ríos G, Koncz C, et al (2004) Functional specialization amongst the Arabidopsis Toc159 family of chloroplast protein import receptors. Plant Cell **16**: 2059-2077

Kuroda H, Maliga P (2003) The plastid *clpP1* protease gene is essential for plant development. Nature **425**: 86-89

Laibach F (1943) *Arabidopsis thaliana* (L.) Heynh. Als Objekt für genetische und entwicklungsphysiologische Untersuchungen. Bot Arch **44**: 439-455

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, et al (2007) Clustal W and Clustal X version 2.0. Bioinformatics **23**: 2947-2948

Lee S, Lee DW, Lee Y, Mayer U, Stierhof YD, Lee S, Jürgens G, Hwang I (2009) Heat shock protein cognate 70-4 and an E3 ubiquitin ligase, CHIP, mediate plastid-destined precursor degradation through the ubiquitin-26S proteasome system in *Arabidopsis*. Plant Cell **21**: 3984-4001

Lenhard JM (2011) Lipogenic enzymes as therapeutic targets for obesity and diabetes. Curr Pharm Des **17**: 325-331

Leutwiler LS, Houghevans BR, Meyerowitz EM (1984) The DNA of *Arabidopsis thaliana*. Mol Gen Genet **194**: 15-23

Li C, Shen Y, Meeley R, McCarty DR, Tan BC (2015) *Embryo defective 14* encodes a plastid-targeted cGTPase essential for embryogenesis in maize. Plant J **84**: 785-799

Li C, Qian W, Maclean CJ, Zhang J (2016) The fitness landscape of a tRNA gene. Science **352**: 837-840

Li HM, Teng YS (2013) Transit peptide design and plastid import regulation. Trends Plant Sci **18**: 360-366

Li X, Ilarslan H, Brachova L, Qian HR, Li L, Che P, Wurtele ES, Nikolau BJ (2011) Reverse-genetic analysis of the two biotin-containing subunit genes of the heteromeric acetyl-coenzyme A carboxylase in Arabidopsis indicates a unidirectional functional redundancy. Plant Physiol **155**: 293-314

Liu TJ, Zhang CY, Yan HF, Zhang L, Ge XJ, Hao G (2016) Complete plastid genome sequence of *Primula sinensis* (Primulaceae): structure comparison, sequence variation and evidence for accD transfer to nucleus. PeerJ **4**: e2101

Liu W, Harrison DK, Chalupska D, Gornicki P, O'Donnell CC, Adkins SW, Haselkorn R, Williams RR (2007) Single-site mutations in the carboxyltransferase domain of plastid acetyl-CoA carboxylase confer resistance to grass-specific herbicides. Proc Natl Acad Sci USA **104**: 3627-3632

Lloyd J, Meinke D (2012) A comprehensive dataset of genes with a loss-of-function mutant phenotype in *Arabidopsis thaliana*. Plant Physiol **158**: 1115-1129

Lü S, Zhao H, Parsons EP, Xu C, Kosma DK, Xu X, Chao D, Lohrey G, Bangarusamy DK, Wang G, et al (2011) The *glossyhead1* allele of *ACC1* reveals a principal role for multidomain acetyl-coenzyme A carboxylase in the biosynthesis of cuticular waxes by Arabidopsis. Plant Physiol **157**: 1079-1092

Lukowitz W, Gillmor CS, Sheible WR (2000) Positional cloning in Arabidopsis: why it feels good to have a genome initiative working for you. Plant Physiol **123**: 795-805

Lyons E, Pedersen B, Kane J, Alam M, Ming R, Tang H, Wang X, Bowers J, Paterson A, Lisch D, et al (2008) Finding and comparing syntenic regions among Arabidopsis and the outgroups papaya, poplar, and grape: CoGe with rosids. Plant Physiol **148**: 1772-1781

Ma Z, Dooner HK (2004) A mutation in the nuclear-encoded plastid ribosomal protein S9 leads to early embryo lethality in maize. Plant J **37**: 92-103

Magnard JL, Heckel T, Massonneau A, Wisniewski JP, Cordelier S, Lassagne H, Perez P, Dumas C, Rogowsky PM (2004) Morphogenesis of maize embryos requires *ZmPRPL35-1* encoding a plastid ribosomal protein. Plant Physiol **134**: 649-663

Maier RM, Neckermann K, Igloi GL, Kossel H (1995) Complete sequence of the maize chloroplast genome: gene content, hotspots of divergence and fine tuning of genetic information by transcript editing. J Mol Biol **251**: 614-628

Mankin AS (1997) Pactamycin resistance mutations in functional sites of 16 S rRNA. J Mol Biol **274**: 8-15

May T, Soll J (2000) 14-3-3 proteins form a guidance complex with chloroplast precursor proteins in plants. Plant Cell **12**: 53-63

McElver J, Tzafrir I, Aux G, Rogers R, Ashby C, Smith K, Thomas C, Schetter A, Zhou Q, Cushman MA, et al (2001) Insertional mutagenesis of genes required for seed development in *Arabidopsis thaliana*. Genetics **159**: 1751-1763

McKhann HI, Camilleri C, Bérard A, Bataillon T, David JL, Reboud X, Le Corre V, Caloustian C, Gut IG, Brunel D (2004) Nested core collections maximizing genetic diversity in *Arabidopsis thaliana*. Plant J **38**: 193–202

Meinke DW, Sussex IM (1979a) Embryo-lethal mutants of *Arabidopsis thaliana*: a model system for genetic analysis of plant embryo development. Devel Biol **72:** 50-61

Meinke DW, Sussex IM (1979b) Isolation and characterization of six embryo-lethal mutants of *Arabidopsis thaliana*. Devel Biol **72:** 62-72

Meinke DW (1985) Embryo-lethal mutants of *Arabidopsis thaliana*: Analysis of mutants with a wide range of lethal phases. Theor Appl Genet **69**: 543-552

Meinke D, Cherry JM, Dean C, Rounsley SD, Koornneef M (1998) *Arabidopsis thaliana*: a model plant for genome analysis. Science **282**: 662-682

Meinke DW, Meinke LK, Showalter TC, Schissel AM, Mueller LA, Tzafrir I (2003) A sequence-based map of Arabidopsis genes with mutant phenotypes. Plant Physiol **131**: 409-418

Meinke D, Muralla R, Sweeney C, Dickerman A (2008) Identifying essential genes in *Arabidopsis thaliana*. Trends Plant Sci **13**: 483-491

Meinke D, Sweeney C, Muralla R (2009) Integrating the genetic and physical maps of *Arabidopsis thaliana*: Identification of mapped alleles of cloned essential (*EMB*) genes. PLoS One **4**: e7386

Meinke D (2013) Large-scale mutant analysis of seed development in *Arabidopsis thaliana*. In: Becraft PW (ed) Seed Genomics, pp 5-20. Wiley-Blackwell, Ames, Iowa

Menninger JR, Coleman RA (1993) Lincosamide antibiotics stimulate dissociation of peptidyl-tRNA from ribosomes. Antimicrob Agents Chemother **37**: 2027-2029

Meyerowitz EM, Pruitt RE (1985) *Arabidopsis thaliana* and plant molecular genetics. Science **229**: 1214-1218

Mochizuki N, Brusslan JA, Larkin R, Nagatani A, Chory J (2001) Arabidopsis genomes uncoupled 5 (GUN5) mutant reveals the involvement of Mg-chelatase H subunit in plastid-to-nucleus signal transduction. Proc Natl Acad Sci USA **98**: 2053-2058

Moreau A, Yotov WV, Glorieux FH, St-Arnaud R (1998) Bone-specific expression of the alpha chain of the nascent polypeptide-associated complex, a coactivator potentiating c-Jun-mediated transcription. Mol Cell Biol **18**: 1312-1321

Motamayor JC, Mockaitis K, Schmutz J, Haiminen N, Livingstone D III, Cornejo O, Findley SD, Zheng P, Utro F, Royaert S, et al (2013) The genome sequence of the most widely cultivated cacao type and its use to identify candidate genes regulating pod color. Genome Biol **14**: r53

Muralla R, Lloyd J, Meinke D (2011) Molecular foundations of reproductive lethality in *Arabidopsis thaliana*. PLoS One **6**: 28398

Naver H, Boudreau E, Rochaix JD (2001) Functional studies of Ycf3: its roles in assembly of photosystem I and interactions with some of its subunits. Plant Cell **13**: 2731-2745

Nobusawa T, Okushima Y, Nagata N, Kojima M, Sakakibara H, Umeda M (1995) Synthesis of very-long-chain fatty acids in the epidermis controls plant organ growth by restricting cell proliferation. PLoS One **11**: e1001531

Nordberg M, Hu TT, Ishino Y, Jhaveri J, Toomajian C, Zheng H, Bakker E, Calabrese P, Gladstone J, Goyal R, et al. (2005) The pattern of polymorphism in *Arabidopsis thaliana*. PLoS Biol **3**: e196

Odintsova MS, Yurina NP (2005) Genomics and evolution of cellular organelles. Genetika **41**: 957-967

Ohlrogge J, Browse J (1995) Lipid biosynthesis. Plant Cell **7**: 957-970

Olmstead RG, Palmer JD (1994) Chloroplast DNA systematics: a review of methods and data analysis. Am J Bot **81**: 1205-1224

Palmer JD, Thompson WF (1982) Chloroplast DNA rearrangements are more frequent when a large inverted repeat sequence is lost. Cell **29**: 537-550

Parker N, Wang Y, Meinke D (2014) Natural variation in sensitivity to a loss of chloroplast translation in Arabidopsis. Plant Physiol **166**: 2013-2027

Parker N, Wang Y, Meinke D (2016) Analysis of Arabidopsis accessions hypersensitive to a loss of chloroplast translation. Plant Physiol **172**: 1862-1875

Patel R, Hsu SC, Bédard J, Inoue K, Jarvis P (2008) The Omp85-related chloroplast outer envelope protein OEP80 is essential for viability in Arabidopsis. Plant Physiol **148**: 235-245

Peske F, Savelsbergh A, Katunin VI, Rodnina MV, Wintermeyer W (2004) Conformational changes of the small ribosomal subunit during elongation factor G-dependent tRNA-mRNA translocation. J Mol Biol **343**: 1183-1194

Ponce-Rojas J, Avendaño-Monsalve M, Yañez-Falcón A, Jaimes-Miranda F, Garay E, Torres-Quiroz F, DeLuna A, Funes S (2017) αβ′-NAC cooperates with Sam37 to mediate early stages of mitochondrial protein import. FEBS J **284**: 814-830

Provart NJ, Alonso J, Assmann SM, Bergmann D, Brady SM, Brkljacic J, Browse J, Chapple C, Colot V, Cutler S, et al (2015) 50 years of Arabidopsis research: highlights and future directions. New Phytol **209**: 921-944

Qin YM, Hu CY, Pang Y, Kastaniotis AJ, Hiltunen JK, Zhu YX (2007) Saturated very-long-chain fatty acids promote cotton fiber and *Arabidopsis* cell elongation by activating ethylene biosynthesis. Plant Cell **19**: 3692-3704

Raffaele S, Vailleau F, Léger A, Joubès J, Miersch O, Huard C, Blée E, Mongrand S, Domergue F, Roby D (2008) A MYB transcription factor regulates very-long-chain fatty acid biosynthesis for activation of the hypersensitive cell death response in *Arabidopsis*. Plant Cell **20**: 752-767

Ramos-Vega M, Guevara-García A, Llamas E, Sánchez-León N, Olmedo-Monfil V, Vielle-Calzada JP, León P (2015) Functional analysis of the *Arabidopsis thaliana CHLOROPLAST BIOGENESIS 19* pentatricopeptide repeat editing protein. New Phytol **208**: 430-441

Rappleye CA, Tagawa A, Le Bot N, Ahringer J, Aroian RV (2003) Involvement of fatty acid pathways and cortical interaction of the pronuclear complex in *Caenorhabditis elegans* embryonic polarity. BMC Dev Biol **3**: 8

Raubeson LA, Jansen RK (1992) Chloroplast DNA evidence on the ancient evolutionary split in vascular land plants. Science **255**: 1697-1699

Rédei GP (1975) Arabidopsis as a genetic tool. Annu Rev Genet **9**: 111-127

Rhee SY, Beavis W, Berardini TZ, Chen G, Dixon D, Doyle A, Garcia-Hernandez M, Huala E, Lander G, Montoya M, et al. (2003) The Arabidopsis Information Resource (TAIR): a model organism database providing a centralized, curated gateway to Arabidopsis biology, research materials and community. Nucleic Acids Res **31**: 224-228

Romani I, Tadini L, Rossi F, Masiero S, Pribil M, Jahns P, Kater M, Leister D, Pesaresi P (2012) Versatile roles of Arabidopsis plastid ribosomal proteins in plant growth and development. Plant J **72**: 922-934

Rousseau-Gueutin M, Huang X, Higginson E, Ayliffe M, Day A, Timmis JN (2013) Potential Functional Replacement of the Plastidic Acetyl-CoA Carboxylase Subunit (*accD*) Gene by Recent Transfers to the Nucleus in Some Angiosperm Lineages. Plant Physiol **161**: 1918-1929

Salie MJ, Thelen JJ (2016) Regulation and structure of the heteromeric acetyl-CoA carboxylase. Biochim Biophys Acta **1861**: 1207-1213

Sasamura T, Matsuno K, Fortini ME (2013) Disruption of *Drosophila melanogaster* lipid metabolism genes causes tissue overgrowth associated with altered developmental signaling. PLoS Genet **9**: e1003917

Sato S, Nakamura Y, Kaneko T, Asamizu E, Tabata S (1999) Complete structure of the chloroplast genome of *Arabidopsis thaliana*. DNA Res **6**: 283-290

Schmitz-Linneweber C, Small I (2008) Pentatricopeptide repeat proteins: a socket set for organelle gene expression. Trends Plant Sci **13**: 663-670

Schneiter R, Guerra CE, Lampl M, Tatzer V, Zellnig G, Klein HL, Kohlwein SD (2000) A novel cold-sensitive allele of the rate-limiting enzyme of fatty acid synthesis, acetyl coenzyme A carboxylase, affects the morphology of the yeast vacuole through acylation of Vac8p. Mol Cell Biol **20**: 2984-2995

Schneiter R, Hitomi M, Ivessa AS, Fasch EV, Kohlwein SD, Tartakoff AM (1996) A yeast acetyl coenzyme A carboxylase mutant links very-long-chain fatty acid synthesis to the structure and function of the nuclear membrane-pore complex. Mol Cell Biol **16**: 7161-7172

Schulte W, Töpfer R, Stracke R, Schell J, Martini N (1997) Multi-functional acetyl-CoA carboxylase from *Brassica napus* is encoded by a multi-gene family: Indication for plastidic localization of at least one isoform. Proc Natl Acad Sci USA **94**: 3465-3470

Shang B, Xu C, Zhang X, Cao H, Xin W, Hu Y (2016) Very-long-chain fatty acids restrict regeneration capacity by confining pericycle competence for callus formation in Arabidopsis. Proc Natl Acad Sci USA **113**: 5101-5106

Shen Y, Li C, McCarty DR, Meeley R, Tan BC (2013) *Embryo defective12* encodes the plastid initiation factor 3 and is essential for embryogenesis in maize. Plant J **74**: 792-804

Shi LX, Theg SM (2013) The chloroplast protein import system: from algae to trees. Biochim Biophys Acta **1833**: 314-331

Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Yamaguchi-Shinozaki K, et al. (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J **5**: 2043-2049

Siemenroth A, Wollgiehn R, Neumann D, Börner T (1981) Synthesis of ribosomal RNA in ribosome-deficient plastids of the mutant "albostrians" of *Hordeum vulgare* L. Planta **153**:

547-555

Slotte T, Hazzouri KM, Ågren JA, Koenig D, Maumus F, Guo YL Steige K, Platts AE, Escobar JS, Newman LK, et al (2013) The *Capsella rubella* genome and the genomic consequences of rapid mating system evolution. Nat Genet **45**: 831-835

Sommerville C, Koornneef M (2002) A fortunate choice: the history of Arabidopsis as a model plant. Nat Rev Genet **3**: 883-889

Sosso D, Canut M, Gendrot G, Dedieu A, Chambrier P, Barkan A, Consonni G, Rogowsky PM (2012) *PPR8522* encodes a chloroplast-targeted pentatricopeptide repeat protein necessary for maize embryogenesis and vegetative development. J Exp Bot **63**: 5843-5857

Stengel A, Benz JP, Buchanan BB, Soll J, Bölter B (2009) Preprotein import into chloroplasts via the Toc and Tic complexes is regulated by redox signals in *Pisum sativum*. Mol Plant **2**: 1181-1197

Straub SCK, Fishbein M, Livshultx T, Foster Z, Parks M, Weitemier K, Cronn RC, Liston A (2011) Building a model: developing genomic resources for common milkweed (*Asclepias syriaca*) with low coverage genome sequencing. BMC Genomic **12**: 211

Sun Y, Moore MJ, Zhang S, Soltis PS, Soltis DE, Zhao T, Meng A, Li X, Li J, Wang H (2016) Phylogenomic and structural analyses of 18 complete plastomes across nearly all families of early-diverging eudicots, including an angiosperm-wide analysis of IR gene content evolution. Mol Phylogenet Evol **96**: 93-101

Swiatek M, Kuras R, Sokolenko A, Higgs D, Olive J, Cinque G, Müller B, Eichacker LA, Stern DB, Bassi R (2001) The chloroplast gene *ycf9* encodes a photosystem II (PSII) core subunit, PsbZ, that participates in PSII supramolecular architecture. Plant Cell **13**: 1347-1367

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol **30**: 2725-2729

Tang L, Yao A, Yuan M, Tang Y, Liu J, Liu X, Qiu R (2016) Transcriptional up-regulation of genes involved in photosynthesis of the Zn/Cd hyperaccumulator *Sedum alfredii* in response to zinc and cadmium. Chemosphere **164**: 190-200

Teng YS, Su YS, Chen LJ, Lee YJ, Hwang I, Li HM (2006) Tic21 is an essential translocon component for protein translocation across the chloroplast inner envelope membrane. Plant Cell **18**: 2247-2257

Tenson T, Lovmar M, Ehrenberg M (2003) The mechanism of action of macrolides, lincosamides and streptogramin B reveals the nascent peptide exit path in the ribosome. J Mol Biol **330**: 1005-1014

Terry MJ, Smith AG (2013) A model for tetrapyrrole synthesis as the primary mechanism for plastid-to-nucleus signaling during chloroplast biogenesis. Front Plant Sci **4**: 14

The 1001 Genomes Consortium (2016) 1,135 genomes reveal the global pattern of polymorphism in *Arabidopsis thaliana*. Cell **166**: 481-491

Tiller N, Bock R (2014) The translational apparatus of plastids and its role in plant development. Mol Plant **7**: 1105-1120

Tong L (2013) Structure and function of biotin-dependent carboxylases. Cell Mol Life Sci **70**: 863-891

Trösch R, Jarvis P (2011) The stromal processing peptidase of chloroplasts is essential in *Arabidopsis*, with knockout mutations causing embryo arrest after the 16-cell stage. PLoS One **6**: e23039

Uchoi A, Malik SK, Choudhary R, Kumar S, Rohini MR, Pal D, Ercisli S, Chaudhury R (2016) Inferring Phylogenetic Relationships of Indian Citron (*Citrus medica* L.) based on *rbcL* and *matK* Sequences of Chloroplast DNA. Biochem Genet **54**: 249-269

Vogel J, Hübschmann T, Börner T, Hess WR (1997) Splicing and intron-internal RNA editing of *trnK-matK* transcripts in Barley plastids: support for MatK as an essential splice factor. J Mol Biol **270**: 179-187

Walbot V, Coe EH (1979) Nuclear gene *iojap* conditions a programmed change to ribosome-less plastids in *Zea mays*. Proc Natl Acad Sci USA **76**: 2760-2764

Walker CJ, Willows RD (1997) Mechanism and regulation of Mg-chelatase. Biochem J **327**: 321-333

Wang Y, Chang H, Hu S, Lu X, Yuan C, Zhang C, Wang P, Xiao L, Xue GP, Guo X (2014) Plastid casein kinase 2 knockout reduces abscisic acid (ABA) sensitivity, thermotolerance, and expression of ABA- and heat-stress-responsive nuclear genes. J Exp Bot **65**: 4159-4175

Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ (2009) Jalview version 2: a multiple sequence alignment editor and analysis workbench. Bioinformatics **25**: 1189-1191

Weatherly SC, Volrath SL, Elich TD (2004) Expression and characterization of recombinant fungal acetyl-CoA carboxylase and isolation of a soraphen-binding domain. Biochem **380**: 105-110

Wei J, Tong L (2015) Crystal structure of the 500-kDa yeast acetyl-CoA carboxylase holoenzyme dimer. Nature **526**: 723-727

Weigel D, Meyerowitz EM (1994) The ABCs of floral homeotic genes. Cell **78**: 203-209

Weigel D (2012) Natural variation in Arabidopsis: from molecular genetics to ecological genomics. Plant Physiol **158**: 2-22

Wirmer J, Westhof E (2006) Molecular contacts between antibiotics and the 30S ribosomal particle. Methods Enzymol **415**: 180-202

Wolfe KH, Li WH, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc Natl Acad Sci USA **84**: 9054-9058

Yang KS, Kim HS, Jin UH, Lee SS, Park JA, Lim YP, Pai HS (2007) Silencing of NbBTF3 results in developmental defects and disturbed gene expression in chloroplasts and mitochondria of higher plants. Planta **225**: 1459-1469

Yu Q, Lutz KA, Maliga P (2017) Efficient Plastid Transformation in Arabidopsis. Plant Physiol **175**: 186-193

Zhang H, Tweel B, Tong L (2004) Molecular basis for the inhibition of the carboxyltransferase domain of acetyl-coenzyme-A carboxylase by haloxyfop and diclofop. Proc Natl Acad Sci USA **101**: 5910-5915

Zhang YF, Hou MM, Tan BC (2013) The requirement of WHIRLY1 for embryogenesis is dependent on genetic background in maize. PLoS One **8**: e67369

Zhou J, Chen X, Cui Y, Sun W, Li Y, Wang Y, Song J, Yao H (2017) Molecular Structure and Phylogenetic Analyses of Complete Chloroplast Genomes of Two *Aristolochia* Medicinal Species. Int J Mol Sci **18**: 1839

Zu X, Zhong J, Luo D, Tan J, Zhang Q, Wu Y, Liu J, Cao R, Wen G, Cao D (2013) Chemical genetics of acetyl-CoA carboxylases. Molecules **18**: 1704-1719

Zubko MK, Day A (1998) Stable albinism induced without mutagenesis: A model for ribosome-free plastid inheritance. Plant J **15**: 265-271

APPENDIX A: Arabidopsis Natural Accessions Analyzed

This appendix lists all 252 natural accessions of Arabidopsis that have been used for spectinomycin analyses in this project. Included data are accession names, seed stock numbers from the Arabidopsis Biological Resource Center (ABRC), whether the seed stock is progeny from a sibling plant to the one sequenced for the 1001 Genomes Project, information on stratification, vernalization and germination problems, reported country of origin, the purpose of the accession for this project, and whether progeny seed stocks were harvested in our lab. Adapted from Parker et al. (2016).

Footnotes for the title row of the following table are described below:

[a] Stratification (S), extended treatment at 4° C used for germination of seeds on plates. Vernalization (V), treatment at 4° C for 5-6 weeks of plants at the rosette stage. Seeds repeated had problems germinating (G).

[b] DEL, Predicted small deletion or frameshift; FS, First forward genetic screen; LUS, Like Unknown Sensitive; NON, Predicted nonsense mutation; RAR, Predicted rearrangement or major deletion; REV-1, Reverse genetic screen – ACC1 conserved; REV-2, Reverse genetic screen – ACC2 conserved; SPL, Predicted splicing defect; SS, Second forward genetic screen; TIC, Reverse genetic screen – TIC20-IV conserved; TRP, Predicted transit peptide variant.

Accession	ABRC Stock Number	1001 Genome Sibling Seed	Growth Information [a]	Reported Country of Origin	Initial Purpose [b]	Secondary Purpose [b]	Progeny Seed Harvested
"Nossen"	Lab				FS	NON	Yes
Aa-0	CS76428	Yes	S	Germany	SS	REV-2	
Ag-0	CS76430	Yes	S	France	SS		
Aitba-1	CS76649	Yes	S / G	Morocco	SS	REV-2	Yes
An-1	CS28015			Belgium	FS		
Ang-0	CS76436	Yes		Belgium	SS		
App1-14	CS76668	Yes		Sweden	REV-1	REV-2	
App1-16	CS76669	Yes	V	Sweden	REV-2		
ARGE-1-15	CS76672	Yes		France	SS		
Ba-1	CS76441	Yes	V	United Kingdom	LUS		
Baa-1	CS76442	Yes	S	Netherlands	SS	REV-2	
Balan-1	CS76687	Yes	V	Russia	REV-2		
Bay-0	CS28056			Germany	FS	REV-2	
Bch-4	CS28060			Germany	FS		
Bd-0	CS76445	Yes	S	Germany	TRP		Yes
Be-1	CS28063			Germany	FS		Yes
Ber	CS76448	Yes		Denmark	SS	REV-2	
Berkeley	CS28067			USA (CA)	FS		
Bik-1	CS76449	Yes	S	Lebanon	SS		
Bil-5	CS76709	Yes	V	Sweden	RAR		
Bl-1	CS76450	Yes		Italy	SS	REV-2	
Bla-1/12	CS28086			Spain	FS		
Blh-1	CS28089			Czech Republic	NON	REV-2	
Blh-1(2)	CS76098			Czech Republic	NON	REV-2	
Boot-1	CS76452	Yes	S	United Kingdom	SS	REV-2	
Bor-4	CS76454	Yes	S	Czech Republic	SS		
Borky1	CS76453	Yes		Czech Republic	REV-2		
BRI-2	CS76725	Yes		France	REV-2		
Bs-1	CS76456	Yes		Switzerland	SS		
Bsch-0	CS76457	Yes		Germany	TRP		
Buckhorn Pass	CS76733	Yes	V	USA (CA)	LUS	REV-2	
Bur-0	CS76734	Yes		Ireland	SS		
C24	CS28127				FS	REV-2	
Cal-0	CS76460	Yes		United Kingdom	TIC		
Can-0	CS76740	Yes		Spain	SS	REV-2	
CATS-6	CS76760	Yes	S / G	France	SPL	REV-2	
Chat-1	CS76463	Yes	S / G	France	SS	REV-2	
Chi-0	CS76464	Yes		Russia	TRP	REV-2	

CIBC-5	CS76465	Yes		United Kingdom	SS		
Co-1	CS76468	Yes		Portugal	SS		
Col-0	Lab				FS	REV-2	Yes
Com-1	CS76469	Yes		France	SS		Yes
CON-7	CS76781	Yes		France	SS	REV-2	
Cvi-0	CS28197			Cape Verde Islands	FS		
CYR	CS76790	Yes		France	SS	REV-2	
Da(1)-12	CS76470	Yes		Czech Republic	SS		
Db-1	CS28203			Germany	FS	REV-2	
Del-10	CS76397	Yes	S	Yugoslavia	SS		
Dem-4	CS76794	Yes	V	USA	LUS	REV-2	
Di-G	CS76472	Yes		France	TRP		Yes
Dja-1	CS76473	Yes	S	Kyrgyzstan	SS	REV-2	
Dog-4	CS76386	Yes	V	Turkey	TRP	REV-2	
Dra3-1	CS76811	Yes	V	Sweden	LUS	REV-2	
Draha2	CS76812	Yes		Czech Republic	SS		
DraIV-6-22	CS76823	Yes		Czech Republic	TIC		
Durh-1	CS76477	Yes		United Kingdom	SS		
Ema-1	CS76480	Yes	S	United Kingdom	SS	REV-2	
En-1	CS28233			Germany	FS		
En-D	CS28230			Ukraine	FS		
Erg2-6	CS76845	Yes		Germany	SS		
Eri-1	CS28240			Sweden	FS		
Est	CS76485	Yes		Germany	SS	REV-2	
Est-0/1	CS28243			Russia	FS		
Etna-2	CS76487	Yes	S / V	Italy	SS	REV-2	Yes
Faneromnemi-3	CS76853	Yes		Greece	SS		
Fei-0	CS28250			Portugal	FS	REV-2	
Fell1-10	CS76855	Yes		Germany	TRP		
Filet-1	CS76858	Yes		Italy	SS		
Ga-0	CS76490	Yes		Germany	SS	REV-2	
Gd-1	CS28275			Germany	FS		
Geg-14	CS76876	Yes		Armenia	SS		
Gel-1	CS76492	Yes		Netherlands	SS		
Giffo-1	CS76878	Yes		Italy	REV-2		
Gifu-2	CS76494	Yes		Japan	SS	REV-2	
Gn-1	CS76880	Yes		Germany	SPL	REV-2	Yes
Gn2-3	CS76881	Yes		Germany	SS		Yes
Go-0	CS28282			Germany	FS		Yes
Gr-1 (Graz)	CS76496	Yes		Austria	SS		
Gradi-1	CS76887	Yes		Croatia	NON		

Gre-0	CS76497	Yes		USA (MI)	LUS	REV-2	
Grivo-1	CS76888	Yes	S / V	Bulgaria	REV-2		Yes
Gu-0	CS28331			Germany	FS		
Gy-0	CS76499	Yes		France	SS	REV-2	
Hag-2	CS76907	Yes	V	Sweden	TIC		
Ha-HBT1-2	CS76898	Yes		Germany	SS		Yes
Ha-HBT2-10	CS76899	Yes		Germany	SS		
Ha-S-B	CS76903	Yes		Germany	SS		
Hi-0	CS28346			Netherlands	FS	REV-2	
Hn-0	CS76513	Yes		Germany	TRP		
Hod	CS76924	Yes		Czech Republic	NON	REV-2	Yes
Hof-1	CS76925	Yes	S	Germany	SPL		
Hovdala-2	CS76937	Yes	V	Sweden	RAR		
HR-10	CS28355			United Kingdom	FS		
Hs-0	CS76515	Yes		Germany	SS		
Hsm	CS76941	Yes		Czech Republic	SS	REV-2	
Iasi-1	CS76944	Yes		Romania	REV-2		
In-0	CS76516	Yes		Austria	SS		
IP-Alo-0	CS76662	Yes	S	Portugal	DEL		
IP-Ber-0	CS78887	Yes	S	Spain	SPL		
IP-Cor-0	CS76782	Yes	S / V	Spain	LUS		
IP-Cum-1	CS76787	Yes	S / (V)	Spain	TRP		Yes
IP-Deh-1	CS76793	Yes	S / V	Spain	TIC		Yes
IP-Gua-1	CS76894	Yes	S	Spain	LUS		
IP-Hom-4	CS76929	Yes	S / V	Spain	LUS		
IP-Lso-0	CS77055	Yes	S	Spain	REV-2		
IP-Mar-1	CS77068	Yes	S	Spain	TIC		
IP-Pal-0	CS77159	Yes	S	Spain	REV-2		
IP-Ren-6	CS77212	Yes	S	Spain	DEL		
IP-Tdc-0	CS77344	Yes	S	Spain	TIC		Yes
IP-Tor-1	CS77378	Yes	S	Spain	REV-2		
IP-Vin-0	CS78846	Yes	S	Spain	DEL		
IP-Vis-0	CS78848	Yes	S	Spain	LUS		
IP-Voz-0	CS78849	Yes	S	Spain	DEL		
Is-0	CS76517	Yes		Germany	TRP		
Jl-3	CS28369			Czech Republic	FS		Yes
Jm-0	CS76520	Yes		Czech Republic	SS		
Kar-1	CS76522	Yes	S	Kyrgyzstan	SS		
Karag-2	CS76961	Yes		Russia	SS		
Kas-2	CS76523	Yes	S	India	SS		
Kb-0	CS76524	Yes		Germany	NON		Yes

Kil-0	CS76526	Yes		United Kingdom	SS	REV-2	
Kin-0	CS76527	Yes		USA (MI)	SS	REV-2	
Kl-1	CS28390			Germany	FS		
Kl-5	CS76528	Yes		Germany	NON		Yes
Kn-0	CS28395			Lithuania	FS	REV-2	
Kni-1	CS76970	Yes	V	Sweden	LUS	REV-2	
Knjas-1	CS76971	Yes		Serbia	SS		
Knox-18	CS76530	Yes		USA (IN)	SS	REV-2	Yes
Koch-1	CS76396	Yes		Ukraine	SS	REV-2	
Kolar-1	CS76974	Yes		Bulgaria	SS		
Koln	CS76976	Yes		Germany	SS		
Kolyv-6	CS76980	Yes		Russia	SS		
Kondara	CS76532	Yes	S	Tajikistan	SS		
K-oze-1	CS76957	Yes		Russia	SS		
Kru-3	CS76986	Yes	V	Sweden	TIC		Yes
Kyoto	CS76535	Yes		Japan	SS	REV-2	
Kz-1	CS28427			Kazakhstan	FS		
La-0	CS76538	Yes		Germany	SS		Yes
LDV-18	CS77013	Yes		France	SS	REV-2	
Ler-1	CS28449			Germany	FS	REV-2	Yes
Leska-1-44	CS77030	Yes	V	Bulgaria	REV-2		
Lip-0	CS76542	Yes		Poland	SS	REV-2	
Litva	CS76543	Yes		Lithuania	SS		
Lm-2	CS28473			France	FS	REV-2	
Lo-1	CS28474			Germany	FS		
Lu3-30	CS77057	Yes		Germany	DEL		
Lu4-2	CS77058	Yes		Germany	DEL		
Mdn-1	CS77077	Yes		USA	LUS	REV-2	
Mer-6	CS76414	Yes	S	Spain	SS		
Mh-0	CS76550	Yes		Poland	SS	REV-2	
Mh-1	CS28493			Poland	FS		Yes
MIC-31	CS77082	Yes		USA (MI)	LUS	REV-2	
MNF-Che-2	CS77096	Yes		USA	REV-2		
MNF-Jac-12	CS77097	Yes		USA (MI)	LUS	REV-2	
MNF-Pot-21	CS77099	Yes		USA	LUS	REV-2	
MNF-Pot-75	CS77100	Yes		USA	LUS	REV-2	
Mt-0	CS28502			Libya	FS		Yes
Mv-0	CS76556	Yes		USA (MA)	LUS	REV-2	
Mz-0	CS28506			Germany	FS		Yes
Nc-1	CS76559	Yes		France	SS		
NC-6	CS77124	Yes		USA (NC)	LUS	REV-2	

Nd-0/1	CS28528			Germany	FS		
Nemrut-1	CS76398	Yes	V	Turkey	TRP	REV-2	
Neo-6	CS76560	Yes		Tajikistan	SS		
Nfa-8	CS28532			United Kingdom	FS	REV-2	
Nie1-2	CS76402			Germany	FS	REV-2	Yes
Nok-3	CS76562	Yes		Netherlands	SS		
Np-0	CS76563	Yes		Germany	SS		
Nw-0	CS76564	Yes		Germany	TRP		
Nz-1	CS28578			New Zealand	FS		
Ob-0	CS76566	Yes		Germany	RAR		Yes
Old-1	CS76567	Yes		Germany	RAR		Yes
Olympia-2	CS77144	Yes		Greece	DEL		
Oy-0	CS28591			Norway	FS	REV-2	Yes
Pa-2	CS28595			Italy	FS		
Ped-0	CS76415	Yes	V	Spain	LUS		
Per-1	CS76571	Yes		Russia	SS		
Pi-0	CS76572	Yes		Austria	SS	REV-2	
Pi-2	CS28639			Austria	FS		Yes
Pla-0	CS76573	Yes		Spain	SS	REV-2	
Pna-10	CS76574	Yes		USA (MI)	SS	REV-2	Yes
Pna-17	CS76575	Yes	V	USA (MI)	LUS	REV-2	
Pog-0	CS76576	Yes	S	Canada	SS	REV-2	
Pro-0	CS76577	Yes	S	Spain	SS		
PT2.21	CS77191	Yes		USA (PT)	LUS	REV-2	
Pu2-23	CS76579	Yes	S	Czech Republic	SS		
Qar-8a	CS76581	Yes	V	Lebanon	DEL		Yes
Qui-0	CS76417			North Africa/Spain	FS	REV-2	
Ra-0	CS28665			France	FS		
Ragl-1	CS76583	Yes		United Kingdom	SS	REV-2	
Rennes-1	CS76586	Yes		France	SS	REV-2	
Rev-2	CS77215	Yes	V	Sweden	RAR		
RLD-2	CS28688			Russia	FS		
Rmx-A01	CS76589	Yes		USA (MI)	LUS	REV-2	
Rmx-A180	CS77218	Yes		USA (MI)	LUS	REV-2	
RRS-10	CS76592	Yes		USA (IN)	SS	REV-2	Yes
RRS-7	CS76593	Yes		USA (IN)	TRP		
Rubexhnoe-1	CS76594	Yes	S	Ukraine	SS		
Sapporo-0	CS28724			Japan	FS		
Sav-0	CS28725			Czech Republic	FS		Yes
Schl-7	CS77240	Yes		Germany	REV-2		
Se-0	CS76597	Yes	S	Spain	SS		

Seattle-0	CS76598	Yes		USA (WA)	SS	REV-2	
Sei-0	CS76599	Yes		Italy	SS	REV-2	
Sha	CS28736			Tajikistan	FS	REV-2	Yes
Sha(2)	CS76382			Tajikistan	FS	REV-2	
Si-0	CS76601	Yes		Germany	SS	REV-2	
Slavi-1	CS76419	Yes	S	Bulgaria	SS		
SLSP-31	CS77254	Yes		USA (MI)	LUS	REV-2	
SLSP-35	CS77255	Yes		USA (MI)	LUS	REV-2	
Smolj-1	CS77256	Yes	S	Bulgaria	REV-2		
Sorbo	CS76602	Yes		Tajikistan	SS		
Spr1-2	CS77261	Yes	V	Sweden	LUS	REV-2	
Spro-1	CS77263	Yes	V	Sweden	RAR		
Spro-2	CS77264	Yes	S / V	Sweden	SPL		
Sq-8	CS76604	Yes		United Kingdom	SS	REV-2	
Star-8	CS76400	Yes		Germany	TRP		
Ste-2	CS77274	Yes	S / V	Sweden	SPL		
Ste-3	CS77275	Yes	S / V	Sweden	SPL		
Stw-0	CS76605	Yes		Russia	SS		
T1020	CS77289	Yes	V	Sweden	RAR		
Ta-0	CS76608	Yes		Czech Republic	SS		
TAMM-2	CS76610	Yes	V	Finland	RAR		
Tha-1	CS76611	Yes		Netherlands	SS		
Tol-0	CS76614	Yes		USA (OH)	SS	REV-2	
Ts-1	CS76615	Yes		Spain	REV-1	REV-2	Yes
Tscha-1	CS76616	Yes		Austria	SS	REV-2	
Tsu-0	CS28780			Japan	FS	REV-2	Yes
Tu-0	CS76617	Yes		Italy	SS	REV-2	Yes
Tul-0	CS76618	Yes		USA (MI)	SS	REV-2	Yes
Ty-0	CS76619	Yes		United Kingdom	SS		
Uk-1	CS76620	Yes	S	Germany	SS		
UKSW06-333	CS78813	Yes		United Kingdom	LUS	REV-2	
Ulies-1	CS78815	Yes		Romania	TIC	REV-2	
Ullapool-8	CS78821	Yes	S	United Kingdom	SPL		
Uod-1	CS76621	Yes		Austria	SS	REV-2	
Utrecht	CS76622	Yes		Netherlands	SS		
Vaar2-6	CS78831	Yes	V	Sweden	RAR		
Van-0	CS28796			Canada	FS	REV-2	
Vimmerby	CS78845	Yes	S / V	Sweden	SPL		
Wa-1	CS76626	Yes		Poland	REV-1	REV-2	Yes
WalhaesB4	CS76408	Yes	S	Germany	REV-2		
WAR	CS78853	Yes		USA (RI)	REV-2		

Wei-0	CS28816			Switzerland	FS		Yes
Wil-1	CS28819			Russia	FS		
Wl-0	CS76630	Yes	S	Germany	SPL		Yes
Ws-2	CS28828			Russia	FS	REV-2	
Yeg-1	CS76394	Yes		Armenia	FS		
Yo-0	CS76633	Yes	V	USA (CA)	LUS	REV-2	
Zal-1	CS76634	Yes	S	Kyrgyzstan	SS		
Zdr-1	CS76635	Yes		Czech Republic	SS	REV-2	
Zu-1	CS28847			Switzerland	FS		

APPENDIX B: Detailed Spectinomycin Responses of Natural Accessions and Knockout Lines Analyzed

This appendix lists the spectinomycin responses of all 252 natural accessions of Arabidopsis, and four relevant knockout lines that have been analyzed this project. Included data are accession names, number of seedlings analyzed, the assigned spectinomycin response category, the response score, and the distribution of seedling phenotypes on spectinomycin. Adapted from Parker et al. (2016).

Footnotes for the title row of the following table are described below:

[a] Higher scores reflect increasing levels of tolerance.

[b] Numbers define classes from expanded cotyledons without leaves (1) to extensive rosettes with sizeable leaves (11) as defined in the text. Refer to Figure 3.7 for examples of seedling phenotypes for each class. Green to red color gradiant based on percentage of seedlings within each phenotypic class; Green is 0% and Red is 100%.

Accession	Seedlings Analyzed	Response Category	Response Score [a]	Percentage of Seedlings Analyzed [b]								
				1	2	3	5	6	7	9	10	11
Chat-1	23	Tolerant	9.7	0.0%	0.0%	0.0%	0.0%	8.7%	0.0%	17.4%	56.5%	17.4%
Be-1	341	Tolerant	9.5	0.3%	0.0%	0.3%	0.3%	0.0%	2.3%	37.8%	55.4%	3.6%
Jl-3	352	Tolerant	9.4	0.8%	2.0%	0.6%	1.7%	0.8%	7.4%	21.0%	44.9%	20.8%
Tu-0	84	Tolerant	9.4	0.0%	0.0%	0.0%	0.0%	0.0%	2.4%	57.1%	39.3%	1.2%
Mh-1	20	Tolerant	9.3	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	70.0%	30.0%	0.0%
Ha-HBT1-2	28	Tolerant	9.1	0.0%	0.0%	0.0%	3.6%	0.0%	3.6%	60.7%	32.1%	0.0%
Mz-0	20	Tolerant	9.1	5.0%	0.0%	0.0%	0.0%	0.0%	0.0%	45.0%	50.0%	0.0%
Pi-2	40	Tolerant	9.0	0.0%	0.0%	0.0%	0.0%	0.0%	5.0%	82.5%	12.5%	0.0%
Kl-1	20	Tolerant	9.0	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	#####	0.0%	0.0%
Wei-0	79	Tolerant	9.0	1.3%	1.3%	0.0%	0.0%	1.3%	2.5%	67.1%	26.5%	0.0%
Tsu-0	490	Tolerant	8.8	0.0%	0.4%	0.4%	1.4%	1.0%	13.5%	63.9%	18.8%	0.6%
Pog-0	60	Tolerant	8.6	0.0%	0.0%	0.0%	0.0%	0.0%	23.3%	68.3%	8.4%	0.0%
Fell1-10	27	Tolerant	8.6	0.0%	0.0%	0.0%	3.7%	0.0%	14.8%	77.8%	3.7%	0.0%
Uod-1	78	Tolerant	8.5	0.0%	2.6%	0.0%	0.0%	1.3%	14.1%	78.2%	3.8%	0.0%
En-D	20	Tolerant	8.5	0.0%	0.0%	0.0%	5.0%	0.0%	20.0%	65.0%	10.0%	0.0%
An-1	20	Tolerant	8.4	0.0%	0.0%	0.0%	0.0%	0.0%	30.0%	70.0%	0.0%	0.0%
Mt-0	20	Tolerant	8.4	0.0%	0.0%	5.0%	5.0%	0.0%	5.0%	85.0%	0.0%	0.0%
Lm-2	70	Tolerant	8.3	1.4%	1.4%	0.0%	2.8%	0.0%	22.9%	62.9%	8.6%	0.0%
Erg2-6	27	Tolerant	8.3	3.7%	0.0%	3.7%	0.0%	0.0%	11.1%	81.5%	0.0%	0.0%
Ema-1	41	Tolerant	8.2	0.0%	2.5%	4.9%	7.3%	0.0%	14.6%	51.2%	17.1%	2.4%
Sorbo	28	Tolerant	8.1	3.6%	0.0%	3.6%	0.0%	0.0%	21.4%	71.4%	0.0%	0.0%
Uk-1	27	High Intermediate	8.3	0.0%	0.0%	0.0%	0.0%	0.0%	33.3%	66.7%	0.0%	0.0%
En-1	20	High Intermediate	8.3	0.0%	0.0%	0.0%	0.0%	0.0%	35.0%	65.0%	0.0%	0.0%
Si-0	83	High Intermediate	8.2	2.4%	0.0%	0.0%	0.0%	1.2%	33.7%	51.8%	9.7%	1.2%
Sha	74	High Intermediate	8.1	0.0%	0.0%	0.0%	5.4%	0.0%	37.8%	50.0%	6.8%	0.0%
Ang-0	28	High Intermediate	8.1	0.0%	0.0%	0.0%	0.0%	0.0%	46.4%	53.6%	0.0%	0.0%
Giffo-1	25	High Intermediate	8.0	0.0%	0.0%	0.0%	8.0%	12.0%	24.0%	36.0%	20.0%	0.0%
C24	73	High Intermediate	7.9	0.0%	2.8%	0.0%	6.8%	2.8%	34.2%	43.8%	9.6%	0.0%
Baa-1	28	High Intermediate	7.8	3.6%	0.0%	0.0%	7.1%	0.0%	35.7%	46.5%	7.1%	0.0%
CYR	76	High Intermediate	7.8	0.0%	2.6%	0.0%	13.2%	2.6%	26.3%	47.4%	7.9%	0.0%
Ag-0	28	High Intermediate	7.6	7.1%	0.0%	7.1%	0.0%	0.0%	25.0%	46.5%	14.3%	0.0%
Blh-1(2)	20	High Intermediate	7.5	10.0%	0.0%	0.0%	10.0%	0.0%	25.0%	40.0%	15.0%	0.0%
Nc-1	20	High Intermediate	7.3	10.0%	0.0%	0.0%	0.0%	10.0%	30.0%	50.0%	0.0%	0.0%
Nz-1	20	High Intermediate	6.4	25.0%	5.0%	0.0%	5.0%	5.0%	10.0%	20.0%	30.0%	0.0%
Yeg-1	20	Intermediate	7.8	5.0%	0.0%	5.0%	0.0%	0.0%	45.0%	10.0%	35.0%	0.0%
Filet-1	27	Intermediate	7.6	0.0%	0.0%	0.0%	3.7%	0.0%	63.0%	29.6%	3.7%	0.0%
Slavi-1	27	Intermediate	7.6	0.0%	0.0%	0.0%	0.0%	3.7%	66.7%	29.6%	0.0%	0.0%
Gd-1	20	Intermediate	7.6	0.0%	0.0%	0.0%	5.0%	20.0%	35.0%	35.0%	5.0%	0.0%
Sha(2)	20	Intermediate	7.6	0.0%	0.0%	0.0%	0.0%	0.0%	75.0%	20.0%	5.0%	0.0%

IP-Lso-0	55	Intermediate	7.5	0.0%	0.0%	7.3%	21.8%	1.8%	20.0%	21.8%	27.3%	0.0%
Lo-1	20	Intermediate	7.5	0.0%	0.0%	0.0%	0.0%	5.0%	70.0%	20.0%	5.0%	0.0%
Bor-4	26	Intermediate	7.4	3.8%	3.8%	0.0%	11.5%	0.0%	34.7%	34.7%	11.5%	0.0%
Aa-0	27	Intermediate	7.3	0.0%	3.7%	3.7%	3.7%	0.0%	55.6%	25.9%	7.4%	0.0%
Co-1	28	Intermediate	7.3	0.0%	0.0%	0.0%	10.7%	3.6%	57.1%	28.6%	0.0%	0.0%
Se-0	28	Intermediate	7.3	0.0%	3.6%	7.1%	0.0%	7.1%	42.9%	35.7%	3.6%	0.0%
ARGE-1-15	28	Intermediate	7.2	0.0%	3.6%	0.0%	3.6%	0.0%	67.8%	25.0%	0.0%	0.0%
Kni-1	56	Intermediate	7.2	1.8%	5.4%	3.6%	10.7%	5.4%	28.5%	32.1%	12.5%	0.0%
Stw-0	28	Intermediate	7.2	0.0%	0.0%	0.0%	0.0%	0.0%	89.3%	10.7%	0.0%	0.0%
Is-0	28	Intermediate	7.2	0.0%	0.0%	0.0%	14.3%	17.9%	35.7%	32.1%	0.0%	0.0%
Bsch-0	24	Intermediate	7.2	0.0%	4.2%	4.2%	4.2%	0.0%	62.4%	12.5%	12.5%	0.0%
Db-1	75	Intermediate	7.2	0.0%	8.0%	2.7%	8.0%	5.3%	32.0%	44.0%	0.0%	0.0%
Karag-2	28	Intermediate	7.1	3.6%	0.0%	0.0%	10.7%	7.1%	46.5%	32.1%	0.0%	0.0%
Nfa-8	20	Intermediate	7.1	5.0%	0.0%	0.0%	0.0%	35.0%	30.0%	20.0%	10.0%	0.0%
Nw-0	28	Intermediate	7.0	0.0%	0.0%	0.0%	3.6%	0.0%	92.8%	3.6%	0.0%	0.0%
Ga-0	28	Intermediate	7.0	3.6%	0.0%	0.0%	14.3%	0.0%	57.1%	25.0%	0.0%	0.0%
Ta-0	28	Intermediate	7.0	0.0%	0.0%	0.0%	3.6%	17.8%	67.9%	10.7%	0.0%	0.0%
Pi-0	28	Intermediate	6.9	3.6%	0.0%	0.0%	0.0%	0.0%	89.3%	7.1%	0.0%	0.0%
Sei-0	56	Intermediate	6.9	0.0%	1.8%	0.0%	21.4%	1.8%	55.4%	19.6%	0.0%	0.0%
CON-7	28	Intermediate	6.8	0.0%	7.1%	0.0%	3.6%	0.0%	78.6%	7.1%	3.6%	0.0%
Borky1	53	Intermediate	6.8	0.0%	0.0%	1.9%	30.2%	0.0%	43.4%	24.5%	0.0%	0.0%
Draha2	27	Intermediate	6.7	3.7%	3.7%	0.0%	7.4%	0.0%	70.4%	14.8%	0.0%	0.0%
Kyoto	28	Intermediate	6.7	0.0%	0.0%	0.0%	17.8%	0.0%	78.6%	3.6%	0.0%	0.0%
Kondara	28	Intermediate	6.7	0.0%	7.1%	0.0%	10.7%	10.7%	53.6%	17.9%	0.0%	0.0%
Vaar2-6	54	Intermediate	6.7	5.5%	3.7%	3.7%	11.1%	16.7%	27.8%	24.1%	5.5%	1.9%
Sq-8	26	Intermediate	6.7	7.7%	0.0%	0.0%	15.4%	7.7%	50.0%	7.7%	11.5%	0.0%
Rev-2	53	Intermediate	6.6	0.0%	0.0%	0.0%	24.5%	1.9%	66.1%	7.5%	0.0%	0.0%
Zu-1	20	Intermediate	6.6	0.0%	0.0%	0.0%	15.0%	10.0%	75.0%	0.0%	0.0%	0.0%
Koln	28	Intermediate	6.6	0.0%	0.0%	3.6%	14.3%	0.0%	82.1%	0.0%	0.0%	0.0%
Zal-1	28	Intermediate	6.5	0.0%	0.0%	0.0%	32.1%	0.0%	60.8%	7.1%	0.0%	0.0%
CATS-6	47	Intermediate	6.4	12.8%	4.3%	2.1%	6.4%	0.0%	46.8%	25.5%	2.1%	0.0%
T1020	49	Intermediate	6.4	0.0%	8.2%	0.0%	28.6%	2.0%	44.9%	12.2%	4.1%	0.0%
Pna-17	52	Intermediate	6.4	0.0%	7.7%	9.6%	19.3%	1.9%	34.6%	26.9%	0.0%	0.0%
Bs-1	28	Intermediate	6.4	0.0%	0.0%	0.0%	46.4%	0.0%	39.3%	14.3%	0.0%	0.0%
Ullapool-8	51	Intermediate	6.3	0.0%	1.9%	0.0%	41.2%	5.9%	35.3%	15.7%	0.0%	0.0%
IP-Gua-1	81	Intermediate	6.3	9.9%	9.9%	0.0%	7.4%	13.6%	28.4%	29.6%	1.2%	0.0%
Balan-1	52	Intermediate	6.2	3.9%	0.0%	1.9%	34.6%	0.0%	48.1%	11.5%	0.0%	0.0%
Mh-0	28	Intermediate	6.2	3.6%	0.0%	0.0%	28.6%	0.0%	67.8%	0.0%	0.0%	0.0%
Lip-0	53	Intermediate	6.2	0.0%	5.7%	0.0%	24.5%	9.4%	56.6%	3.8%	0.0%	0.0%
SLSP-35	50	Intermediate	6.2	2.0%	4.0%	8.0%	18.0%	4.0%	54.0%	10.0%	0.0%	0.0%
IP-Pal-0	51	Intermediate	6.1	2.0%	7.8%	3.9%	9.8%	21.6%	45.1%	7.8%	2.0%	0.0%

Koch-1	25	Intermediate	6.1	4.0%	0.0%	0.0%	32.0%	16.0%	40.0%	8.0%	0.0%	0.0%
Hsm	24	Intermediate	6.1	0.0%	0.0%	0.0%	45.8%	0.0%	54.2%	0.0%	0.0%	0.0%
Kas-2	26	Intermediate	6.1	7.7%	3.8%	3.8%	23.1%	3.8%	38.5%	19.3%	0.0%	0.0%
Seattle-0	28	Intermediate	6.1	3.6%	10.7%	7.1%	3.6%	3.6%	60.7%	10.7%	0.0%	0.0%
Pro-0	28	Intermediate	6.0	14.3%	7.1%	7.1%	7.1%	3.6%	25.0%	35.8%	0.0%	0.0%
Spr1-2	47	Intermediate	6.0	0.0%	6.4%	12.8%	25.5%	6.4%	27.7%	19.1%	2.1%	0.0%
Da(1)-12	24	Intermediate	6.0	4.2%	0.0%	4.2%	20.8%	25.0%	41.6%	4.2%	0.0%	0.0%
Bil-5	46	Intermediate	6.0	6.5%	2.2%	8.7%	15.2%	4.4%	54.3%	8.7%	0.0%	0.0%
Rubexhnoe-1	27	Intermediate	6.0	3.7%	3.7%	0.0%	37.0%	3.7%	44.4%	7.5%	0.0%	0.0%
Kin-0	23	Intermediate	6.0	8.7%	13.0%	0.0%	4.4%	4.4%	56.5%	13.0%	0.0%	0.0%
Eri-1	20	Intermediate	6.0	5.0%	0.0%	5.0%	20.0%	35.0%	25.0%	10.0%	0.0%	0.0%
Kn-0	19	Intermediate	5.9	0.0%	5.3%	0.0%	26.3%	26.3%	42.1%	0.0%	0.0%	0.0%
Hof-1	56	Intermediate	5.9	0.0%	0.0%	3.6%	48.2%	10.7%	33.9%	3.6%	0.0%	0.0%
Zdr-1	28	Intermediate	5.8	0.0%	0.0%	0.0%	60.7%	7.1%	28.6%	3.6%	0.0%	0.0%
Ha-HBT2-10	27	Intermediate	5.8	0.0%	14.8%	7.4%	3.7%	25.9%	40.8%	7.4%	0.0%	0.0%
toc34 (ppi3-2)	18	Intermediate	5.8	0.0%	0.0%	27.8%	27.8%	0.0%	22.2%	22.2%	0.0%	0.0%
Hn-0	28	Intermediate	5.7	0.0%	3.6%	0.0%	53.6%	0.0%	42.8%	0.0%	0.0%	0.0%
Hi-0	20	Intermediate	5.7	0.0%	0.0%	5.0%	30.0%	55.0%	10.0%	0.0%	0.0%	0.0%
App1-16	54	Intermediate	5.6	0.0%	0.0%	0.0%	66.7%	1.8%	31.5%	0.0%	0.0%	0.0%
Col-0	287	Intermediate	5.6	1.8%	8.7%	9.1%	38.3%	1.8%	25.4%	12.5%	2.4%	0.0%
WalhaesB4	39	Intermediate	5.6	15.4%	5.1%	2.6%	23.1%	0.0%	35.9%	17.9%	0.0%	0.0%
Geg-14	28	Intermediate	5.5	3.6%	7.1%	0.0%	10.7%	67.9%	10.7%	0.0%	0.0%	0.0%
IP-Voz-0	72	Intermediate	5.5	1.4%	0.0%	4.2%	62.5%	1.4%	29.1%	1.4%	0.0%	0.0%
Pu2-23	23	Intermediate	5.4	4.4%	4.4%	13.0%	21.7%	39.1%	4.4%	13.0%	0.0%	0.0%
Leska-1-44	52	Intermediate	5.4	0.0%	7.7%	7.7%	50.0%	1.9%	26.9%	5.8%	0.0%	0.0%
Tscha-1	28	Intermediate	5.4	10.7%	0.0%	0.0%	50.0%	7.1%	28.6%	0.0%	3.6%	0.0%
Van-0	20	Intermediate	5.4	0.0%	10.0%	15.0%	45.0%	5.0%	5.0%	20.0%	0.0%	0.0%
Hod	72	Intermediate	5.3	1.4%	1.4%	4.1%	66.7%	0.0%	26.4%	0.0%	0.0%	0.0%
Com-1	32	Intermediate	5.3	6.3%	21.9%	15.5%	6.3%	0.0%	25.0%	21.9%	3.1%	0.0%
Cvi-0	20	Intermediate	5.3	5.0%	20.0%	0.0%	0.0%	40.0%	35.0%	0.0%	0.0%	0.0%
Dra3-1	56	Intermediate	5.3	1.8%	10.7%	17.9%	23.2%	10.7%	25.0%	10.7%	0.0%	0.0%
Wa-1	40	Intermediate	5.3	7.5%	10.0%	2.5%	40.0%	0.0%	35.0%	5.0%	0.0%	0.0%
Ws-2	20	Intermediate	5.3	0.0%	0.0%	15.0%	55.0%	5.0%	25.0%	0.0%	0.0%	0.0%
Hs-0	25	Intermediate	5.2	0.0%	0.0%	20.0%	60.0%	0.0%	8.0%	12.0%	0.0%	0.0%
Boot-1	26	Intermediate	5.2	3.8%	3.8%	11.5%	42.4%	3.8%	34.7%	0.0%	0.0%	0.0%
Kar-1	28	Intermediate	5.2	32.1%	7.1%	0.0%	10.7%	0.0%	14.3%	35.8%	0.0%	0.0%
Chi-0	75	Intermediate	5.1	0.0%	1.3%	9.4%	73.3%	0.0%	16.0%	0.0%	0.0%	0.0%
Nie1-2	235	Intermediate	5.1	5.5%	16.6%	6.0%	17.9%	22.5%	28.1%	3.4%	0.0%	0.0%
UKSW06-333	15	Intermediate	5.1	13.3%	13.3%	6.7%	26.7%	6.7%	13.3%	20.0%	0.0%	0.0%
BRI-2	31	Intermediate	5.1	0.0%	3.2%	9.7%	64.5%	9.7%	12.9%	0.0%	0.0%	0.0%
Est	27	Intermediate	5.0	0.0%	18.5%	7.4%	0.0%	74.1%	0.0%	0.0%	0.0%	0.0%

Iasi-1	44	Intermediate	5.0	2.3%	15.9%	11.3%	20.5%	31.8%	13.6%	2.3%	2.3%	0.0%
Jm-0	27	Intermediate	5.0	11.1%	3.7%	0.0%	55.6%	3.7%	25.9%	0.0%	0.0%	0.0%
Fei-0	20	Intermediate	4.9	25.0%	15.0%	0.0%	5.0%	25.0%	10.0%	10.0%	10.0%	0.0%
Kz-1	20	Intermediate	4.9	10.0%	5.0%	15.0%	25.0%	15.0%	30.0%	0.0%	0.0%	0.0%
Np-0	28	Intermediate	4.8	3.6%	7.1%	10.7%	57.2%	10.7%	7.1%	3.6%	0.0%	0.0%
Gel-1	28	Intermediate	4.8	10.7%	0.0%	3.6%	67.8%	3.6%	14.3%	0.0%	0.0%	0.0%
Schl-7	70	Intermediate	4.8	1.4%	15.7%	11.4%	37.2%	11.4%	22.9%	0.0%	0.0%	0.0%
Gr-1 (Graz)	26	Intermediate	4.8	0.0%	11.5%	11.5%	57.7%	0.0%	19.3%	0.0%	0.0%	0.0%
Ba-1	22	Intermediate	4.8	4.5%	18.3%	18.3%	22.7%	4.5%	22.7%	4.5%	4.5%	0.0%
K-oze-1	28	Intermediate	4.8	7.1%	17.9%	7.1%	28.6%	3.6%	35.7%	0.0%	0.0%	0.0%
In-0	84	Intermediate	4.8	13.1%	8.3%	10.6%	31.0%	4.8%	28.6%	3.6%	0.0%	0.0%
Can-0	96	Intermediate	4.7	1.0%	5.2%	30.2%	38.6%	1.0%	21.9%	2.1%	0.0%	0.0%
RRS-7	28	Intermediate	4.7	7.1%	25.0%	3.6%	21.4%	3.6%	39.3%	0.0%	0.0%	0.0%
Durh-1	17	Intermediate	4.7	35.3%	11.8%	0.0%	0.0%	5.9%	23.5%	23.5%	0.0%	0.0%
Star-8	27	Intermediate	4.7	7.4%	29.7%	11.1%	14.8%	0.0%	18.5%	18.5%	0.0%	0.0%
Gy-0	26	Intermediate	4.6	15.4%	19.3%	0.0%	19.3%	15.4%	26.8%	3.8%	0.0%	0.0%
Ha-S-B	28	Intermediate	4.6	3.6%	14.3%	17.9%	35.7%	0.0%	28.5%	0.0%	0.0%	0.0%
Dja-1	52	Intermediate	4.6	3.8%	23.1%	3.8%	40.4%	11.6%	15.4%	1.9%	0.0%	0.0%
Nd-0/1	19	Intermediate	4.6	15.8%	21.1%	5.2%	0.0%	21.1%	36.8%	0.0%	0.0%	0.0%
Gu-0	20	Intermediate	4.6	0.0%	5.0%	25.0%	60.0%	0.0%	10.0%	0.0%	0.0%	0.0%
Hovdala-2	56	Intermediate	4.5	0.0%	0.0%	30.4%	62.5%	0.0%	7.1%	0.0%	0.0%	0.0%
IP-Hom-4	94	Intermediate	4.5	6.4%	22.3%	17.0%	16.0%	13.8%	17.0%	5.4%	2.1%	0.0%
Smolj-1	51	Intermediate	4.5	17.6%	25.5%	0.0%	7.8%	21.6%	15.7%	11.8%	0.0%	0.0%
Kolar-1	27	Intermediate	4.5	29.7%	11.1%	3.7%	11.1%	0.0%	33.3%	11.1%	0.0%	0.0%
Ler-1	20	Intermediate	4.5	10.0%	25.0%	0.0%	0.0%	65.0%	0.0%	0.0%	0.0%	0.0%
Nok-3	26	Intermediate	4.5	30.8%	15.4%	0.0%	3.8%	3.8%	34.7%	11.5%	0.0%	0.0%
Neo-6	27	Intermediate	4.5	18.6%	25.9%	3.7%	11.1%	3.7%	25.9%	3.7%	7.4%	0.0%
Ber	25	Intermediate	4.4	0.0%	4.0%	32.0%	52.0%	4.0%	8.0%	0.0%	0.0%	0.0%
CIBC-5	27	Intermediate	4.4	14.9%	7.4%	22.2%	22.2%	7.4%	22.2%	3.7%	0.0%	0.0%
Ragl-1	27	Intermediate	4.3	14.8%	14.8%	7.4%	29.7%	14.8%	18.5%	0.0%	0.0%	0.0%
MNF-Che-2	53	Intermediate	4.3	0.0%	5.7%	39.6%	41.5%	1.9%	11.3%	0.0%	0.0%	0.0%
Pla-0	27	Intermediate	4.3	11.1%	26.0%	3.7%	22.2%	18.5%	18.5%	0.0%	0.0%	0.0%
Del-10	24	Intermediate	4.2	12.5%	8.3%	16.7%	50.0%	0.0%	12.5%	0.0%	0.0%	0.0%
Nemrut-1	75	Intermediate	4.2	5.3%	38.7%	0.0%	2.7%	53.3%	0.0%	0.0%	0.0%	0.0%
RLD-2	19	Intermediate	4.1	5.3%	36.8%	0.0%	15.8%	42.1%	0.0%	0.0%	0.0%	0.0%
Kolyv-6	27	Intermediate	4.0	18.5%	18.5%	11.1%	29.7%	11.1%	3.7%	7.4%	0.0%	0.0%
Litva	27	Intermediate	3.9	7.4%	3.7%	37.0%	48.2%	0.0%	3.7%	0.0%	0.0%	0.0%
Spro-1	52	Low Intermediate	5.1	3.8%	17.3%	36.6%	7.7%	0.0%	3.8%	13.5%	13.5%	3.8%
Mer-6	28	Low Intermediate	4.5	0.0%	42.9%	14.3%	7.1%	10.7%	3.6%	21.4%	0.0%	0.0%
Yo-0	51	Low Intermediate	4.5	2.0%	35.3%	17.6%	9.8%	0.0%	23.5%	11.8%	0.0%	0.0%
IP-Tor-1	46	Low Intermediate	4.1	4.3%	30.4%	19.6%	17.4%	10.9%	13.1%	4.3%	0.0%	0.0%

Kil-0	28	Low Intermediate	4.0	0.0%	3.6%	46.4%	42.9%	7.1%	0.0%	0.0%	0.0%	0.0%
Gradi-1	8	Low Intermediate	4.0	37.5%	12.5%	0.0%	0.0%	12.5%	37.5%	0.0%	0.0%	0.0%
Ulies-1	109	Low Intermediate	3.9	7.3%	27.5%	23.0%	12.8%	20.2%	2.8%	6.4%	0.0%	0.0%
Rennes-1	44	Low Intermediate	3.9	13.6%	27.3%	18.2%	15.9%	2.3%	15.9%	6.8%	0.0%	0.0%
Bla-1/12	20	Low Intermediate	3.9	20.0%	30.0%	5.0%	0.0%	35.0%	5.0%	5.0%	0.0%	0.0%
Bik-1	28	Low Intermediate	3.8	0.0%	39.3%	17.8%	28.6%	0.0%	10.7%	3.6%	0.0%	0.0%
LDV-18	21	Low Intermediate	3.7	9.5%	33.4%	19.0%	9.5%	19.0%	4.8%	4.8%	0.0%	0.0%
App1-14	38	Low Intermediate	3.7	7.9%	31.6%	13.1%	31.6%	7.9%	7.9%	0.0%	0.0%	0.0%
Berkeley	20	Low Intermediate	3.7	10.0%	30.0%	10.0%	35.0%	10.0%	5.0%	0.0%	0.0%	0.0%
Dog-4	53	Low Intermediate	3.7	26.4%	32.1%	5.7%	1.9%	11.3%	9.4%	13.2%	0.0%	0.0%
WAR	39	Low Intermediate	3.7	5.2%	20.5%	33.3%	25.6%	12.8%	2.6%	0.0%	0.0%	0.0%
Utrecht	28	Low Intermediate	3.6	0.0%	39.3%	21.4%	17.9%	21.4%	0.0%	0.0%	0.0%	0.0%
Hag-2	73	Low Intermediate	3.6	6.9%	37.0%	13.7%	19.2%	20.5%	2.7%	0.0%	0.0%	0.0%
Bay-0	68	Low Intermediate	3.6	13.2%	42.7%	1.5%	8.8%	25.0%	8.8%	0.0%	0.0%	0.0%
DraIV-6-22	79	Low Intermediate	3.5	7.6%	30.4%	29.1%	16.5%	5.1%	8.9%	2.4%	0.0%	0.0%
Tha-1	84	Low Intermediate	3.4	17.9%	39.3%	7.1%	6.0%	21.4%	7.1%	1.2%	0.0%	0.0%
Knjas-1	22	Low Intermediate	3.4	27.3%	31.8%	9.1%	9.1%	4.5%	9.1%	9.1%	0.0%	0.0%
IP-Vis-0	47	Low Intermediate	3.3	31.9%	19.1%	12.8%	14.9%	6.4%	14.9%	0.0%	0.0%	0.0%
Sapporo-0	16	Low Intermediate	3.2	18.8%	37.5%	0.0%	31.2%	12.5%	0.0%	0.0%	0.0%	0.0%
Ra-0	20	Low Intermediate	3.2	40.0%	15.0%	0.0%	20.0%	25.0%	0.0%	0.0%	0.0%	0.0%
Kru-3	173	Sensitive	4.1	7.0%	30.6%	21.4%	10.4%	15.0%	5.8%	5.8%	4.0%	0.0%
IP-Tdc-0	130	Sensitive	3.8	9.2%	23.8%	28.5%	18.5%	4.6%	9.2%	5.4%	0.8%	0.0%
IP-Deh-1	94	Sensitive	3.6	3.2%	24.5%	44.7%	12.8%	1.0%	10.6%	3.2%	0.0%	0.0%
IP-Mar-1	51	Sensitive	3.4	11.8%	17.6%	51.0%	3.9%	0.0%	9.8%	5.9%	0.0%	0.0%
Ty-0	28	Sensitive	3.2	0.0%	57.1%	17.9%	0.0%	25.0%	0.0%	0.0%	0.0%	0.0%
Bch-4	20	Sensitive	3.2	10.0%	45.0%	15.0%	10.0%	20.0%	0.0%	0.0%	0.0%	0.0%
Ped-0	56	Sensitive	3.1	26.8%	17.9%	28.6%	19.6%	0.0%	0.0%	7.1%	0.0%	0.0%
Go-0	58	Sensitive	3.1	17.2%	32.8%	20.7%	15.5%	12.1%	1.7%	0.0%	0.0%	0.0%
Mdn-1	43	Sensitive	2.9	37.2%	16.3%	20.9%	7.0%	4.6%	14.0%	0.0%	0.0%	0.0%
IP-Ren-6	42	Sensitive	2.8	2.4%	57.1%	23.8%	9.5%	4.8%	2.4%	0.0%	0.0%	0.0%
Aitba-1	53	Sensitive	2.8	39.6%	30.2%	5.7%	7.5%	3.8%	9.4%	1.9%	1.9%	0.0%
TAMM-2	54	Sensitive	2.8	0.0%	55.6%	33.3%	5.5%	1.9%	3.7%	0.0%	0.0%	0.0%
IP-Cor-0	8	Sensitive	2.6	25.0%	50.0%	0.0%	12.5%	12.5%	0.0%	0.0%	0.0%	0.0%
Qui-0	18	Sensitive	2.5	44.4%	33.3%	0.0%	0.0%	16.7%	5.6%	0.0%	0.0%	0.0%
MNF-Jac-12	33	Sensitive	2.5	42.5%	12.1%	30.3%	3.0%	9.1%	3.0%	0.0%	0.0%	0.0%
Di-G	82	Sensitive	2.5	12.2%	61.0%	14.6%	3.7%	8.5%	0.0%	0.0%	0.0%	0.0%
Pa-2	20	Sensitive	2.5	10.0%	75.0%	0.0%	5.0%	10.0%	0.0%	0.0%	0.0%	0.0%
Ts-1	76	Sensitive	2.4	21.1%	61.8%	3.9%	2.6%	5.3%	4.0%	1.3%	0.0%	0.0%
Faneromnemi-3	27	Sensitive	2.4	29.6%	29.6%	26.0%	14.8%	0.0%	0.0%	0.0%	0.0%	0.0%
Wil-1	20	Sensitive	2.4	0.0%	90.0%	0.0%	0.0%	10.0%	0.0%	0.0%	0.0%	0.0%
Bl-1	49	Sensitive	2.4	6.1%	59.3%	30.6%	2.0%	2.0%	0.0%	0.0%	0.0%	0.0%

Per-1	27	Sensitive	2.4	18.5%	59.3%	7.4%	11.1%	3.7%	0.0%	0.0%	0.0%	0.0%
Est-0/1	20	Sensitive	2.4	5.0%	85.0%	0.0%	0.0%	10.0%	0.0%	0.0%	0.0%	0.0%
HR-10	20	Sensitive	2.4	50.0%	25.0%	5.0%	0.0%	20.0%	0.0%	0.0%	0.0%	0.0%
Cal-0	112	Sensitive	2.3	8.0%	59.8%	29.5%	0.9%	1.8%	0.0%	0.0%	0.0%	0.0%
IP-Cum-1	129	Sensitive	2.3	34.9%	20.1%	35.7%	9.3%	0.0%	0.0%	0.0%	0.0%	0.0%
“Nossen”	571	Sensitive	2.3	35.0%	23.5%	33.5%	5.4%	0.7%	1.4%	0.5%	0.0%	0.0%
Bur-0	70	Sensitive	2.3	1.4%	75.7%	21.5%	0.0%	1.4%	0.0%	0.0%	0.0%	0.0%
Qar-8a	175	Sensitive	2.2	19.4%	53.2%	21.7%	5.7%	0.0%	0.0%	0.0%	0.0%	0.0%
MNF-Pot-75	32	Sensitive	2.2	31.3%	25.0%	40.6%	3.1%	0.0%	0.0%	0.0%	0.0%	0.0%
Bd-0	79	Sensitive	2.2	7.6%	72.1%	19.0%	0.0%	1.3%	0.0%	0.0%	0.0%	0.0%
MIC-31	49	Sensitive	2.1	32.7%	20.4%	46.9%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Oy-0	229	Sensitive	2.1	9.6%	74.7%	13.5%	0.9%	0.9%	0.4%	0.0%	0.0%	0.0%
Rmx-A01	36	Sensitive	2.1	27.8%	36.1%	36.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Grivo-1	73	Sensitive	2.0	24.7%	53.4%	20.5%	0.0%	1.4%	0.0%	0.0%	0.0%	0.0%
Tol-0	73	Sensitive	2.0	50.8%	21.9%	21.9%	2.7%	0.0%	0.0%	2.7%	0.0%	0.0%
La-0	106	Sensitive	2.0	16.0%	74.5%	8.5%	0.0%	1.0%	0.0%	0.0%	0.0%	0.0%
Olympia-2	33	Sensitive	1.9	30.3%	45.5%	24.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Etna-2	111	Sensitive	1.9	26.1%	64.9%	7.2%	0.9%	0.0%	0.9%	0.0%	0.0%	0.0%
Rmx-A180	42	Sensitive	1.8	42.9%	35.7%	19.0%	2.4%	0.0%	0.0%	0.0%	0.0%	0.0%
Dem-4	28	Sensitive	1.7	57.1%	17.9%	25.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Tul-0	81	Sensitive	1.7	66.7%	11.1%	18.6%	1.2%	1.2%	1.2%	0.0%	0.0%	0.0%
Buckhorn Pass	29	Sensitive	1.7	58.6%	17.3%	24.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
NC-6	54	Sensitive	1.6	50.0%	35.2%	14.8%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
tic20-IV-2 Koncz 11324	9	Sensitive	1.6	88.9%	0.0%	0.0%	0.0%	11.1%	0.0%	0.0%	0.0%	0.0%
PT2.21	55	Sensitive	1.5	56.4%	36.4%	7.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Pna-10	112	Sensitive	1.4	71.4%	15.2%	12.5%	0.9%	0.0%	0.0%	0.0%	0.0%	0.0%
Wl-0	79	Sensitive	1.4	74.6%	12.7%	12.7%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
SLSP-31	53	Sensitive	1.4	86.8%	5.7%	1.9%	3.7%	0.0%	1.9%	0.0%	0.0%	0.0%
Kb-0	73	Sensitive	1.4	74.0%	16.4%	9.6%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Gre-0	48	Sensitive	1.4	79.2%	6.2%	14.6%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Gifu-2	78	Sensitive	1.3	78.2%	14.1%	5.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
RRS-10	81	Sensitive	1.3	85.2%	6.2%	6.2%	1.2%	0.0%	0.0%	1.2%	0.0%	0.0%
Spro-2	80	Sensitive	1.3	85.0%	3.8%	10.0%	1.2%	0.0%	0.0%	0.0%	0.0%	0.0%
IP-Ber-0	49	Sensitive	1.3	81.6%	10.2%	8.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Blh-1	71	Sensitive	1.3	80.3%	14.1%	5.6%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
MNF-Pot-21	50	Hypersensitive	1.4	66.0%	32.0%	2.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Lu4-2	55	Hypersensitive	1.3	80.0%	14.5%	5.5%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Sav-0	275	Hypersensitive	1.2	84.0%	13.1%	2.2%	0.0%	0.7%	0.0%	0.0%	0.0%	0.0%
Ob-0	74	Hypersensitive	1.2	86.5%	8.1%	5.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Old-1	75	Hypersensitive	1.2	84.0%	14.7%	1.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

IP-Vin-0	33	Hypersensitive	1.2	84.8%	15.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Gn2-3	191	Hypersensitive	1.1	89.0%	8.4%	2.6%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Gn-1	83	Hypersensitive	1.1	91.6%	3.6%	4.8%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Mv-0	56	Hypersensitive	1.1	91.1%	5.3%	3.6%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Knox-18	80	Hypersensitive	1.1	91.2%	7.5%	1.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Kl-5	76	Hypersensitive	1.1	94.7%	1.3%	4.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ste-2	83	Hypersensitive	1.1	92.8%	6.0%	1.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Lu3-30	54	Hypersensitive	1.1	92.6%	7.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
IP-Alo-0	51	Hypersensitive	1.1	96.0%	2.0%	2.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
tic20-IV-1 Sail-97-F10	877	Hypersensitive	1.0	97.6%	1.1%	1.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
acc2 Salk_148966C	1218	Hypersensitive	1.0	97.2%	2.1%	0.7%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ste-3	82	Hypersensitive	1.0	97.6%	2.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Vimmerby	67	Hypersensitive	1.0	98.5%	1.5%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

APPENDIX C: Details on Enhancer Phenotype Classes of TT Plants from a Tsu-0 Cross with *emb3126-1*

This appendix lists the details of the progeny plants screened from a single cross (Tsu-0 x *emb3126-1*). All plants listed are homozygous Tsu-0 for the suppressor (*ACC2*). Included data are the phenotype class, identification numbers of the plants screened, who each plant was screened by, plant generation, proposed genotype of the enhancer, genotype of three loci linked to the enhancer (*TOC34*, *EMB3137*, and *OEP80*), whether the plant is a proposed recombinant line, number of embryos measured, the average and range of embryos length in µm, percent of embryos < 100 µm, > 100 µm and >200 µm, and percent of embryo phenotypes globular, triangular, linear and cotyledon.

Footnotes for the title row of the following table are described below:

[a] Late, plants homozygous Tsu-0 for the enhancer. Interm, plants heterozygous for the enhancer. Early, plants homozygous "Nossen" for the enhancer. Parentheses, borderline plants.

[b] DM, David Meinke. NP, Nicole Parker.

[c] T, homozygous Tsu-0. H, heterozygous. N, homozygous "Nossen".

[d] TOCx3137, crossover between *TOC34* and *EMB3137*. 3137xOEP, crossover between *EMB3137* and *OEP80*. NA, recombinant line identified in previous generation.

Phenotype Class [a]	Plants Screened	Screened By [b]	Plant Generation	Proposed Enhancer Genotype [c]	*TOC34 EMB3137 OEP80* Genotype [c]	Proposed Recombinant [d]	Embryos Measured	Embryo Lengths (µm)		Percent Embryos by Length			Percent Embryos by Stage			
								Avg.	Range	<100 µm	>100 µm	>200 µm	Glob.	Triang.	Linear	Cotyl.
Late	1B	DM	F2	T	T - T - T	No	64	194	100 - 400	0.0	98.4	35.9	0.0	25.0	60.9	14.1
Late	4D-2B	DM	F3	T	T - T - T	No	94	182	80 - 450	2.1	94.7	25.5	1.1	33.0	42.5	23.4
Late	7D-3B	DM	F3	T	T - T - T	No	17	178	100 - 270	0.0	100.0	29.4	0.0	47.1	23.5	29.4
Late	S2-10D-2D	DM	F3	T	T - T - T	No	44	171	90 - 320	2.3	93.2	22.7	6.8	13.6	68.2	11.4
Late	19E-2E	NP	F3	T	T - T - T	No	58	171	110 - 310	0.0	100.0	20.7	0.0	24.1	67.3	8.6
Late	20D	DM	F2	T	T - T - T	No	41	169	90 - 500	2.4	90.2	19.5	2.4	34.2	43.9	19.5
Late	4D-1A	DM	F3	T	T - T - T	No	37	169	110 - 430	0.0	100.0	8.1	0.0	35.1	56.8	8.1
Late	7D-2B	DM	F3	T	T - T - T	No	17	166	100-300	0.0	82.4	23.5	17.6	47.1	5.9	29.4
Late	4D-2A	DM	F3	T	T - T - T	No	36	164	100 - 320	0.0	88.9	13.9	8.3	25.0	58.4	8.3
Late	4D-2E	DM	F3	T	T - T - T	No	50	163	100 - 250	0.0	98.0	6.0	0.0	18.0	82.0	0.0
Late	19E-3D	NP	F3	T	T - T - T	No	49	152	110 - 210	0.0	100.0	2.0	0.0	36.7	63.3	0.0
Late	7D-3A	DM	F3	T	T - T - T	No	21	148	110 - 360	0.0	100.0	9.5	0.0	66.7	23.8	9.5
Late	4D-4E	DM	F3	T	H - T - T	TOCx3137	75	146	80 - 240	2.7	96.0	2.7	2.7	50.7	45.3	1.3
Late	4D-4B	DM	F3	T	T - T - H	3137xOEP	40	145	80 - 260	5.0	92.5	7.5	5.0	47.5	40.0	7.5
Late	S2-10D-1E	DM	F3	T	T - T - H	3137xOEP	49	145	90 - 300	4.1	81.6	8.2	6.1	49.0	38.8	6.1
Late	16E	DM+NP	F2	T	H - T - T	TOCx3137	50	142	100 - 290	0.0	90.0	14.0	0.0	84.0	2.0	14.0
Late	3B-2B	DM	F3	T	T - T - T	No	46	142	80 - 260	10.9	82.6	6.5	17.4	67.4	8.7	6.5
Late	3B-2D	DM	F3	T	H - T - T	TOCx3137	39	140	100 - 270	0.0	87.2	12.8	12.8	71.8	2.6	12.8
Late	10A	DM	F2	T	T - T - T	No	20	138	100 - 190	0.0	95.0	0.0	0.0	75.0	25.0	0.0
Late	19E-4E	NP	F3	T	T - T - T	No	49	137	90 - 240	6.1	79.6	4.1	22.4	42.9	32.6	2.1
Late	19E-4A	NP	F3	T	T - T - T	No	44	136	90 - 220	4.5	79.5	6.8	22.7	45.5	31.8	0.0
Late	7D-6A	DM	F3	T	T - T - T	No	29	133	100 - 270	0.0	86.2	6.9	13.8	75.9	3.4	6.9
Late	17B	DM+NP	F2	T	T - T - H	3137xOEP	40	129	100 - 220	0.0	82.5	2.5	7.5	77.5	7.5	7.5
Late	19E-1B	NP	F3	T	H - T - T	TOCx3137	96	126	70 - 200	9.4	74.0	1.0	32.3	41.7	26.0	0.0
Late	S2-10D-3E	DM	F3	T	T - T - H	3137xOEP	64	123	80 - 230	7.8	71.9	3.1	28.1	54.7	15.6	1.6
Late	3B-1A-5D	DM	F4	T	N - T - T	NA	54	119	80 - 250	5.6	77.8	1.9	22.2	75.9	1.9	0.0
Late	7D-4B-2E	DM	F4	T	N - T - T	NA	96	117	80 - 360	11.5	64.6	1.0	17.7	72.9	8.3	1.1
Late	3B-1A-1D	DM	F4	T	N - T - T	NA	51	115	80 - 150	7.8	78.4	0.0	9.8	88.2	2.0	0.0
(Late)	S2-3B-4A	NP	F3	T	T - T - T	NA	94	114	70 - 210	31.9	57.4	1.1	42.6	42.6	14.8	0.0
(Late)	S2-10D-2B	DM	F3	T	T - T - T	No	51	106	50 - 170	37.3	56.9	0.0	41.2	52.9	5.9	0.0

(Late)	3B-1A-3A	DM	F4	T	N - T - T	NA	68	99	70 - 130	33.8	30.9	0.0	70.6	29.4	0.0	0.0
(Late)	S2-3B-2D	DM	F3	T	T - T - T	NA	66	95	60 - 230	54.5	19.7	1.5	83.3	9.1	6.1	1.5
(Interm)	S2-8D	NP	F2	H	H - H - H	No	30	143	90 - 300	6.7	80.0	20.0	30.0	33.3	33.3	3.4
(Interm)	S2-8D-5A	DM	F3	H	H - H - H	No	56	136	70 - 400	28.6	66.1	14.3	32.2	37.5	23.2	7.1
Interm	7D	DM	F2	H	H - H - H	No	34	125	60 - 500	47.1	41.2	5.9	58.8	20.6	20.6	0.0
Interm	S2-8D-5B	DM	F3	H	H - H - H	No	44	120	70 - 400	40.9	31.8	11.4	56.8	25.0	4.5	13.7
Interm	S2-8D-1D	DM	F3	H	H - H - H	No	44	120	50 - 370	34.1	47.7	6.8	45.5	38.6	9.1	6.8
Interm	S2-4A	NP	F2	H	H - H - H	No	16	114	80 - 180	43.8	50.0	0.0	81.2	12.5	6.3	0.0
Interm	4D-5D	DM+NP	F3	H	H - H - H	No Prog.Seed	75	113	50 - 280	53.3	38.7	10.7	60.0	14.7	24.0	1.3
Interm	S2-8D-2D	DM	F3	H	N - H - H	TOCx3137	48	111	60 - 280	45.8	41.7	8.3	58.4	22.9	10.4	8.3
Interm	S2-10E	NP	F2	H	H - H - H	No	22	110	70 - 230	54.5	27.3	9.1	68.2	9.1	22.7	0.0
Interm	S2-8D-5E	DM	F3	H	H - H - H	No	76	107	60 - 320	65.8	31.6	9.2	65.8	19.7	6.6	7.9
Interm	4D-5E	NP	F3	H	H - H - H	No	48	107	60 - 310	60.4	29.2	8.3	70.8	10.4	18.8	0.0
Interm	S2-5B	NP	F2	H	H - H - H	No	22	107	60 - 230	68.2	27.3	9.1	77.3	0.0	22.7	0.0
Interm	3B-2A	DM	F3	H	H - H - T	3137xOEP	26	105	70 - 140	23.1	46.2	0.0	69.2	30.8	0.0	0.0
Interm	20B	NP	F2	H	H - H - H	No	17	104	50 - 350	64.7	29.4	5.9	70.5	11.8	11.8	5.9
Interm	7D-4B	DM	F3	H	N - H - H	TOCx3137	22	104	70 - 170	54.5	36.4	0.0	63.6	31.8	4.6	0.0
Interm	3B-1D	DM	F3	H	H - H - H	No	29	102	60 - 240	48.3	31.0	3.5	69.0	27.6	0.0	3.4
Interm	3B-3A	DM	F3	H	H - H - H	No	26	100	60 - 250	38.5	46.2	3.9	69.2	26.9	0.0	3.9
Interm	4D-2D	DM	F3	H	H - H - H	No	61	99	50 - 270	55.7	31.1	4.9	72.1	11.5	16.4	0.0
Interm	S2-10D-2A	DM	F3	H	H - H - H	No	52	99	50 - 270	59.6	36.5	1.9	63.5	19.2	15.4	1.9
Interm	4D-4A	DM	F3	H	H - H - H	No	56	98	50 - 170	57.1	33.9	0.0	69.6	17.9	12.5	0.0
Interm	7D-3E	DM	F3	H	H - H - N	3137xOEP	18	97	60 - 210	77.8	16.6	5.6	83.2	5.6	5.6	5.6
Interm	3B	DM	F2	H	H - H - H	No	36	96	50 - 260	61.1	27.8	5.6	72.2	22.2	0.0	5.6
Interm	S2-10D-3D	DM	F3	H	H - H - H	No	21	95	60 - 180	61.9	33.3	0.0	66.7	23.8	9.5	0.0
Interm	7D-2D	DM	F3	H	H - H - H	No	48	94	60 - 250	64.6	20.8	4.2	79.2	14.6	2.1	4.1
Interm	S2-4B	NP	F2	H	H - H - H	No	53	92	50 - 220	71.7	20.8	1.9	84.9	3.8	11.3	0.0
Interm	4D-6E	NP	F3	H	H - H - N	3137xOEP	43	91	50 - 190	72.1	18.6	0.0	81.4	9.3	9.3	0.0
Interm	3B-3D	DM	F3	H	H - H - H	No	45	91	50 - 240	71.1	22.2	2.2	77.8	20.0	0.0	2.2
Interm	19E-4D	NP	F3	H	H - H - H	No	56	90	50 - 230	69.6	28.6	1.8	71.4	12.5	14.3	1.8
Interm	4D-6B	NP	F3	H	H - H - H	No Prog.Seed	48	87	50 - 210	72.9	16.7	2.1	83.3	6.3	10.4	0.0
Interm	19E-3E	NP	F3	H	H - H - H	No	49	87	60 - 190	69.4	12.2	0.0	91.8	4.1	4.1	0.0
Interm	17A	DM+NP	F2	H	H - H - T	3137xOEP	27	87	50 - 180	66.7	22.2	0.0	70.4	14.8	14.8	0.0
Interm	3B-1A-4B	DM	F4	H	N - H - H	NA	68	87	50 - 320	72.1	26.5	1.5	70.6	22.1	7.3	0.0
Interm	7D-4B-3A	DM	F4	H	N - H - H	NA	47	86	50 - 170	78.7	19.1	0.0	80.8	12.8	6.4	0.0

Interm	7D-4B-1D	DM	F4	H	N - H - T	(Yes) 3137xOEP	42	86	50 - 140	71.4	21.4	0.0	73.8	26.2	0.0	0.0
Interm	7D-4B-3D	DM	F4	H	N - H - H	NA	50	85	60 - 250	78.0	14.0	2.0	84.0	14.0	0.0	2.0
Interm	S2-10D-3B	DM	F3	H	H - H - H	No	62	85	50 - 160	67.7	19.4	0.0	80.6	11.3	8.1	0.0
Interm	7D-4B-3E	DM	F4	H	N - H - H	NA	45	83	60 - 140	73.3	8.9	0.0	86.7	13.3	0.0	0.0
Interm	3B-1A-4D	DM	F4	H	N - H - T	(Yes) 3137xOEP	47	83	60 - 150	78.7	19.1	0.0	80.9	17.0	2.1	0.0
Interm	3B-1A-3E	DM	F4	H	N - H - H	NA	54	83	60 - 140	70.4	18.5	0.0	79.6	20.4	0.0	0.0
Interm	4D-1D	DM	F3	H	H - H - H	No	61	82	50 - 250	82.0	6.6	1.6	93.4	3.3	3.3	0.0
Interm	19E-1D	NP	F3	H	H - H - H	No	74	82	50 - 180	78.4	14.9	0.0	87.8	4.1	8.1	0.0
Interm	19E-4B	NP	F3	H	H - H - H	No	50	82	50 - 210	80.0	14.0	2.0	84.0	8.0	8.0	0.0
Interm	19E-2A	NP	F3	H	H - H - H	No	54	82	50 - 190	74.1	16.7	0.0	83.3	9.3	7.4	0.0
Interm	7D-4A	DM	F3	H	H - H - H	No	51	82	60 - 200	80.4	13.7	0.0	86.2	9.8	2.0	2.0
Interm	S2-10D-3A	DM	F3	H	H - H - H	No	75	81	50 - 170	85.3	10.7	0.0	89.4	5.3	5.3	0.0
Interm	19E-2B	NP	F3	H	H - H - H	No	41	81	50 - 180	73.2	14.6	0.0	68.3	17.1	14.6	0.0
Interm	S2-3B	NP	F2	H	H - H - H	No	27	81	50 - 130	77.8	11.1	0.0	96.3	3.7	0.0	0.0
Interm	S2-10D	NP	F2	H	H - H - H	No	20	81	60 - 110	75.0	10.0	0.0	95.0	5.0	0.0	0.0
Interm	3B-1A-5E	DM	F4	H	N - H - N	(Yes) 3137xOEP	47	81	50 - 140	70.2	17.0	0.0	78.7	21.3	0.0	0.0
Interm	3B-1A-2B	DM	F4	H	N - H - H	NA	50	80	50 - 140	72.0	18.0	0.0	80.0	20.0	0.0	0.0
Interm	S2-10D-2E	DM	F3	H	N - H - H	TOCx3137	52	80	60 - 140	75.0	11.5	0.0	88.5	11.5	0.0	0.0
Interm	19E	DM	F2	H	H - H - H	No	24	80	50 - 140	75.0	16.7	0.0	83.3	16.7	0.0	0.0
Interm	3B-1A-1B	DM	F4	H	N - H - H	NA	54	79	50 - 150	83.3	9.3	0.0	90.7	7.4	1.9	0.0
Interm	3B-1A-2D	DM	F4	H	N - H - H	NA	53	79	50 - 140	75.5	17.0	0.0	79.2	20.8	0.0	0.0
Interm	7D-4B-2B	DM	F4	H	N - H - T	(Yes) 3137xOEP	40	79	50 - 170	82.5	15.0	0.0	85.0	7.5	7.5	0.0
Interm	4D	DM	F2	H	H - H - H	No	39	77	50 - 130	82.1	7.7	0.0	89.7	10.3	0.0	0.0
Interm	11D	DM	F2	H	H - H - H	No	32	75	50 - 120	81.3	6.3	0.0	87.5	12.5	0.0	0.0
Interm	3B-1A	DM	F3	H	N - H - H	TOCx3137	18	75	60 - 100	88.9	0.0	0.0	100.0	0.0	0.0	0.0
?	3B-1A-3D	DM	F4	?	N - H - H	Intriguing	41	66	50 - 120	95.1	2.4	0.0	97.6	2.4	0.0	0.0
Early	S2-6E	NP	F2	N	N - N - N	No	43	85	60 - 210	83.7	4.7	2.3	97.7	0.0	2.3	0.0
Early	S2-8D-4A	DM	F3	N	N - N - N	No	49	83	60 - 110	81.6	6.1	0.0	93.9	6.1	0.0	0.0
Early	S2-8D-5D	DM	F3	N	N - N - N	No	52	82	70 - 120	78.8	5.8	0.0	94.2	5.8	0.0	0.0
Early	14D	NP	F2	N	N - N - N	No	15	82	50 - 110	66.7	6.7	0.0	100.0	0.0	0.0	0.0
Early	S2-6B	NP	F2	N	N - N - N	No	19	82	60 - 110	78.9	5.3	0.0	100.0	0.0	0.0	0.0
Early	S2-8D-3B	DM	F3	N	N - N - H	3137xOEP	63	77	50 - 110	92.1	1.6	0.0	98.4	1.6	0.0	0.0
Early	3B-3B	DM	F3	N	N - N - N	No	23	77	60 - 90	100.0	0.0	0.0	100.0	0.0	0.0	0.0

Early	S2-8D-4E	DM	F3	N	H - N - N	TOCx3137	54	74	60 - 100	96.3	0.0	0.0	100.0	0.0	0.0	0.0
Early	3B-3E	DM	F3	N	N - N - H	3137xOEP	18	71	50 - 80	100.0	0.0	0.0	100.0	0.0	0.0	0.0
Early	3B-1B	DM	F3	N	H - N - N	TOCx3137	29	70	50 - 100	96.6	0.0	0.0	100.0	0.0	0.0	0.0
Early	4D-5A	NP	F3	N	N - N - N	No Prog.Seed	47	70	50 - 100	97.9	0.0	0.0	100.0	0.0	0.0	0.0
Early	S2-8D-6E	DM	F3	N	H - N - N	TOCx3137	54	69	60 - 100	98.1	0.0	0.0	100.0	0.0	0.0	0.0
Early	S2-7B	NP	F2	N	N - N - N	No	20	69	50 - 80	100.0	0.0	0.0	100.0	0.0	0.0	0.0
Early	S2-3D	NP	F2	N	N - N - N	No	24	67	50 - 90	100.0	0.0	0.0	100.0	0.0	0.0	0.0
Early	4D-3D	DM	F3	N	N - N - N	No	29	67	50 - 90	100.0	0.0	0.0	100.0	0.0	0.0	0.0
Early	19E-1A	NP	F3	N	N - N - N	No	48	67	50 - 100	97.9	0.0	0.0	100.0	0.0	0.0	0.0
Early	1A	DM	F2	N	N - N - H	3137xOEP	42	66	50 - 100	97.6	0.0	0.0	100.0	0.0	0.0	0.0
Early	8B	DM+NP	F2	N	N - N - N	No	39	66	50 - 90	100.0	0.0	0.0	100.0	0.0	0.0	0.0
Early	20A	DM+NP	F2	N	N - N - N	No	25	65	50 - 80	100.0	0.0	0.0	100.0	0.0	0.0	0.0
Early	8D	DM+NP	F2	N	N - N - N	No	28	64	50 - 80	100.0	0.0	0.0	100.0	0.0	0.0	0.0
Early	7D-4B-4A	DM	F4	N	N - N - N	NA	40	64	50 - 80	100.0	0.0	0.0	100.0	0.0	0.0	0.0
Early	S2-3B-6E	NP	F3	N	N - N - N	NA	57	63	50 - 90	100.0	0.0	0.0	100.0	0.0	0.0	0.0
Early	3B-1A-4A	DM	F4	N	N - N - N	NA	67	62	50 - 80	100.0	0.0	0.0	100.0	0.0	0.0	0.0
Early	7D-6D	DM	F3	N	N - N - N	No	35	62	50 - 70	100.0	0.0	0.0	100.0	0.0	0.0	0.0
Early	7E	NP	F2	N	N - N - N	No / Transform	38	61	50 - 100	97.4	0.0	0.0	100.0	0.0	0.0	0.0
Early	S2-10D-4B	DM	F3	N	N - N - N	No	59	61	50 - 80	100.0	0.0	0.0	100.0	0.0	0.0	0.0
Early	S2-10D-4D	DM	F3	N	N - N - N	No	54	60	50 - 80	100.0	0.0	0.0	100.0	0.0	0.0	0.0
Early	12E	DM	F2	N	N - N - N	No	29	58	50 - 70	100.0	0.0	0.0	100.0	0.0	0.0	0.0
Early	3B-1A-1A	DM	F4	N	N - N - N	NA	46	58	50 - 70	100.0	0.0	0.0	100.0	0.0	0.0	0.0
Early	13A	NP	F2	N	N - N - N	No / Transform	29	56	50 - 70	100.0	0.0	0.0	100.0	0.0	0.0	0.0
Early	S2-3B-6B	DM+NP	F3	N	N - N - N	NA	82	56	50 - 70	100.0	0.0	0.0	100.0	0.0	0.0	0.0

APPENDIX D: Details on Modifier Phenotype Classes of TT Plants from a Tsu-0 Cross with *emb3126-1*

This appendix lists the details of the progeny plants screened from a single cross (Tsu-0 x *emb3126-1*). All plants listed are homozygous Tsu-0 for the suppressor (*ACC2*) and the enhancer. Included data are the phenotype class, identification numbers of the plants screened, who each plant was screened by, plant generation, number of embryos measured, the average and range of embryos length in µm, percent of embryos > 100 µm, > 200 µm and > 300 µm, and percent of embryo phenotypes globular, triangular, linear and cotyledon.

Footnotes for the title row of the following table are described below:

[a] Late-Adv, progeny plants from the "Late" class of F2 plants with the highest level of embryo rescue. Late-Mod, progeny plants from the "Late" class of F2 plants with a moderate level of embryo rescue. Late-Red, progeny plants from the "Late" class of F2 plants with the lowest level of embryo rescue. L-A-Late, progeny plants from the "Late-Adv" class of F3 plants with the highest level of embryo rescue. L-A-Mod, progeny plants from the "Late-Adv" class of F3 plants with a moderate level of embryo rescue. Borderline, progeny plants from the "Late-Red" class of F3 plants with a moderate level of embryo rescue. L-R-Red, progeny plants from the "Late-Red" class of F3 plants with the lowest level of embryo rescue.

[b] NP, Nicole Parker.

Phenotype Class [a]	Plants Screened	Screened By [b]	Plant Generation	Embryos Measured	Embryo Lengths (µm)		Percent Embryos by Length			Percent Embryos by Stage			
					Average	Range	>100 µm	>200 µm	>300 µm	Globular	Triangular	Linear	Cotyledon
Late-Adv	1B-3B	NP	F3	44	288	100 - 510	97.7	68.2	43.2	0.0	11.3	36.4	52.3
Late-Adv	20D-S2-1D	NP	F3	86	259	110 - 510	100.0	57.0	39.5	0.0	15.1	38.4	46.5
Late-Adv	20D-3A	NP	F3	39	243	130 - 480	100.0	59.0	17.9	0.0	5.1	53.8	41.1
Late-Adv	20D-1E	NP	F3	76	234	100 - 430	98.7	61.8	19.7	0.0	10.5	54.0	35.5
Late-Mod	1B-6B	NP	F3	31	219	120 - 490	100.0	41.9	16.1	0.0	16.1	71.0	12.9
Late-Mod	20D-3B	NP	F3	36	197	100 - 400	94.4	36.1	13.9	11.1	19.4	22.2	47.3
Late-Mod	1B-6A	NP	F3	31	193	100 - 580	93.5	25.8	6.5	6.4	22.6	58.1	12.9
Late-Mod	20D-2B	NP	F3	33	192	90 - 350	93.9	42.4	6.1	3.0	18.2	24.2	54.6
Late-Mod	20D-2D	NP	F3	37	182	110 - 340	100.0	24.3	2.7	2.7	5.4	81.1	10.8
Late-Mod	1B-2E	NP	F3	34	180	120 - 350	100.0	20.6	2.9	0.0	14.7	79.4	5.9
Late-Mod	20D-S2-1B	NP	F3	80	180	90 - 390	97.5	23.8	7.5	2.5	35.0	48.8	13.8
Late-Mod	20D-S2-2D	NP	F3	77	178	110 - 470	100.0	16.9	3.9	0.0	16.9	75.3	7.8
Late-Mod	20D-2A	NP	F3	31	175	90 - 360	93.5	25.8	6.5	12.9	38.7	12.9	35.5
Late-Mod	1B-1D	NP	F3	29	173	110 - 510	100.0	13.8	6.9	3.4	27.6	55.2	13.8
Late-Mod	1B-6D	NP	F3	33	171	100 - 270	97.0	21.2	0.0	9.1	18.2	57.6	15.1
Late-Mod	1B-1E	NP	F3	22	170	110 - 300	100.0	27.3	0.0	13.6	9.1	77.3	0.0
Late-Red	1B-1B	NP	F3	26	157	100 - 400	96.2	7.7	3.8	3.8	7.7	80.8	7.7
Late-Red	1B-3D	NP	F3	37	157	110 - 230	100.0	5.4	0.0	5.4	27.0	67.6	0.0
Late-Red	20D-1B	NP	F3	77	152	100 - 330	98.7	6.5	1.3	7.8	40.2	44.2	7.8
Late-Red	20D-1D	NP	F3	32	152	110 - 230	100.0	9.4	0.0	12.5	40.6	40.6	6.3
Late-Red	20D-3E	NP	F3	41	145	100 - 230	97.6	4.9	0.0	12.2	43.9	34.1	9.8
Late-Red	20D-S2-1E	NP	F3	74	144	90 - 380	73.0	16.2	1.4	31.1	37.8	20.3	10.8
Late-Red	20D-S2-2A	NP	F3	76	136	90 - 330	78.9	2.6	1.3	15.8	59.2	23.7	1.3
Late-Red	20D-S2-2E	NP	F3	72	125	80 - 190	83.3	0.0	0.0	29.1	54.2	16.7	0.0
L-A-Late	1B-3B-2B	NP	F4	27	358	170 - 540	100.0	88.9	66.7	0.0	0.0	22.2	77.8

L-A-Late	1B-3B-2E	NP	F4	52	350	220 - 590	100.0	100.0	61.5	0.0	0.0	36.5	63.5
L-A-Late	1B-3B-1A	NP	F4	55	328	200 - 540	100.0	94.5	54.5	0.0	0.0	29.1	70.9
L-A-Mod	1B-3B-1B	NP	F4	50	289	170 - 530	100.0	82.0	40.0	0.0	0.0	38.0	62.0
L-A-Mod	1B-3B-2D	NP	F4	31	285	170 - 410	100.0	90.3	38.7	0.0	0.0	38.7	61.3
L-A-Mod	20D-3A-2A	NP	F4	38	274	120 - 450	100.0	78.9	34.2	0.0	5.3	36.8	57.9
L-A-Mod	1B-3B-2A	NP	F4	27	269	160 - 380	100.0	77.8	37.0	0.0	0.0	51.9	48.1
L-A-Mod	20D-3A-2E	NP	F4	100	261	120 - 570	100.0	62.0	33.0	0.0	12.0	39.0	49.0
L-A-Mod	1B-3B-1E	NP	F4	50	254	160 - 560	100.0	68.0	24.0	0.0	0.0	72.0	28.0
L-A-Mod	20D-3A-2D	NP	F4	39	251	120 - 460	100.0	51.3	30.8	0.0	7.6	46.2	46.2
L-A-Mod	1B-3B-1D	NP	F4	59	240	160 - 390	100.0	67.8	16.9	0.0	0.0	69.5	30.5
L-A-Mod	20D-3A-1D	NP	F4	38	234	130 - 430	100.0	50.0	23.7	0.0	5.3	55.3	39.4
L-A-Mod	20D-3A-1E	NP	F4	36	231	120 - 410	100.0	47.2	22.2	0.0	5.6	58.3	36.1
L-A-Mod	20D-3A-1B	NP	F4	101	207	110 - 470	100.0	36.6	10.9	0.0	11.9	76.2	11.9
Borderline	20D-3E-1A	NP	F4	62	174	80 - 380	87.1	22.6	3.2	14.5	22.6	45.2	17.7
Borderline	20D-1D-1A	NP	F4	45	165	90 - 390	84.4	20.0	8.9	20.0	35.5	26.7	17.8
Borderline	20D-1D-2B	NP	F4	47	162	90 - 340	87.2	17.0	6.4	12.8	42.6	25.5	19.1
L-R-Red	20D-1D-1E	NP	F4	41	158	100 - 260	92.7	7.3	0.0	12.2	24.4	56.1	7.3
L-R-Red	20D-1D-2D	NP	F4	56	154	110 - 200	100.0	0.0	0.0	0.0	41.1	58.9	0.0
L-R-Red	20D-3E-1B	NP	F4	54	149	80 - 290	85.2	14.8	0.0	14.8	50.0	20.4	14.8
L-R-Red	20D-3E-3E	NP	F4	55	146	80 - 300	83.6	12.7	0.0	20.0	41.8	23.6	14.6
L-R-Red	20D-1D-1D	NP	F4	54	142	90 - 280	87.0	7.4	0.0	16.7	46.3	33.3	3.7
L-R-Red	20D-1D-2E	NP	F4	50	136	80 - 250	86.0	6.0	0.0	20.0	56.0	16.0	8.0
L-R-Red	20D-3E-1D	NP	F4	67	135	80 - 310	73.1	6.0	1.5	29.8	44.8	19.4	6.0
L-R-Red	20D-3E-2D	NP	F4	70	134	80 - 290	75.7	7.1	0.0	27.1	48.6	20.0	4.3
L-R-Red	20D-1D-2A	NP	F4	42	133	80 - 310	83.3	7.2	2.4	21.4	57.1	14.3	7.2
L-R-Red	20D-3E-3A	NP	F4	49	131	80 - 390	69.4	6.1	2.0	34.7	42.9	22.4	0.0

APPENDIX E: 855 Sequenced Accessions from the 1001 Genomes Project

This appendix lists the names for all 855 Arabidopsis accessions used in the *ACC1* and *ACC2* sequence alignments. All sequence data for these accessions was accessed through the Salk 1001 Genomes Browser (http://signal.salk.edu/atg1001/3.0/gebrowser.php).

11C1; ARGE-1-15; ARR-17; Aa-0; Abd-0; Adam-1; Aedal-1; Aedal-3; Ag-0; Agu-1; Aiell-1; Aitba-1; Ak-1; Alc-0; Ale-Stenar-44-4; Ale-Stenar-56-14; Ale-Stenar-64-24; Algustrum; Alst-1; Alt-1; Altai-5; Amel-1; An-1; Ang-0; Anholt-1; Ann-1; Anz-0; App1-12; App1-14; App1-16; Appt-1; BEZ-9; BI-4; BRE-14; BRI-2; Ba-1; Baa1-2; Baa4-1; Baa5-1; Baa-1; Bach-7; Bach2-1; Bai-10; Bak-2; Bak-7; Balan-1; Basta-1; Basta-2; Bay-0; Bch-1; Bd-0; Bela-1; Bela-2; Benk-1; Ber; Berg-1; Bg-2; Bijisk-4; Bik-1; Bil-5; Bil-7; Bivio-1; Bl-1; Bla-1.7015.MPI; Bla-1.SALK; Blh-1; Boo2-1; Boot-1; Bor-1; Bor-4; Borky1; Br-0; Broesarp-34-145; Broet1-6; Bs-1; Bsch-0; Bu-0; Buckhorn-Pass; Bur-0.MPI; Bur-0.WTC; C24; CATS-6; CHA-41; CIBC-17; CIBC-5; CON-7; CSHL-5; CYR; Ca-0; Cal-0; Can-0; Castelfed-1-197; Castelfed-4-211; Castelfed-4-214; Cdm-0; Cerv-1; Chaba-2; Chat-1; Chi-0; Cimin-1; Cnt-1.5726.MPI; Cnt-1.SALK; Co; Co-1; Col-0; Com-1; Corig-1; Ct-1; Cvi-0.SALK; Cvi-0.SALK; DIR-9; Da1-12; Db-1; Del-10; Dem-4; Di-G; Dja-1; Do-0; Doer-10; Dog-4; Dolen-1; Dolna-1; Don-0; Dospa-1; Doubravnik7; Dr-0; Dra2-1; Dra3-1; DraII-6; DraIII-1; DraII-1; DraIV.5893; DraIV.5907; DraIV.5950; DraIV.5984; DraIV-6-22.5993; Dra-0; Draha2; Duk; Durh-1; ENC-2-1; ESP-1-11; Eden-1; Eden-2; Eden-7; Eden-9; Edi-0; Eds-1; Eds-9; Ei-2; El-0; Ema-1; En-1; En-2; En-D; Epidauros-1; Er-0; Erg2-6; Es-0; Est-1; Est; Et-0; Etna-2; Ey15-2; Faeb-2; Faeb-4; Fael-1; Faneronemi-3; Fei-0; Fell1-10; Fell2-4; Fell3-7; Fi-0; Filet-1; Fjae1-1; Fjae1-2; Fjae1-5; Fjae2-4; Fly2-1; Fly2-2; Fondi-1; Fr-2; Fri-2; Furni-1; GEN-8; Ga-0; Ge-0; Geg-14; Gel-1; Gie-0; Giffo-1; Gifu-2; Gn-1; Gn2-3; Goced-1; Gol-2; Got-22; Got-7; Gr-5; Gr-1; Gradi-1; Gre-0; Grivo-1; Gro-3; Groen-12; Groen-14; Groen-5; Gu-0; Gy-0; HE-1; HKT2; HR-10; HR-5; HSm; Ha-HBT1-2; Ha-HBT2-10; Ha-HBT3-1; Ha-P-13; Ha-P2-1; Ha-S-B; Ha-SP-2; Ha-0; Had-1; Had-2; Haes-1; Hag-2; Hal-1; Ham-1; Hart-2; Hau-0; Hel-3; Hey-1; Hh-0; Hi-0; Hn-0; Hod; Hof-1; Hola-1-1; Hola-2-2; Hola-1-2; Hov1-10; Hov1-7; Hov3-2; Hov3-5; Hov4-1; Hovdala-2; Hs-0; ICE1; ICE102; ICE104; ICE106; ICE107; ICE111; ICE112; ICE119; ICE120; ICE127; ICE130; ICE134; ICE138; ICE150; ICE152; ICE153; ICE163; ICE169; ICE173; ICE181; ICE21; ICE212; ICE213; ICE216; ICE226; ICE228; ICE29; ICE33; ICE36; ICE49; ICE50; ICE60; ICE61; ICE63; ICE7; ICE70; ICE71; ICE72; ICE73; ICE75; ICE79; ICE91; ICE92; ICE93; ICE97; ICE98; IP-Adm-0; IP-Ala-0; IP-All-0; IP-Alm-0; IP-Alo-0; IP-Ang-0; IP-Ara-4; IP-Bar-1; IP-Bea-0; IP-Ben-0; IP-Ber-0; IP-Bis-0; IP-Cab-3; IP-Cad-0; IP-Cal-0; IP-Cap-1; IP-Car-1; IP-Cdc-3;

IP-Cdo-0; IP-Cem-0; IP-Cmo-3; IP-Coa-0; IP-Coc-1; IP-Cor-0; IP-Cum-1; IP-Cur-4; IP-Deh-1; IP-Elb-0; IP-Fue-2; IP-Fun-0; IP-Gra-0; IP-Gua-1; IP-Her-12; IP-Hom-4; IP-Hor-0; IP-Hum-2; IP-Iso-4; IP-Jim-1; IP-Lab-7; IP-Ldd-0; IP-Lso-0; IP-Mar-1; IP-Men-2; IP-Moa-0; IP-Moc-11; IP-Mon-5; IP-Mos-1; IP-Mot-0; IP-Mun-0; IP-Mur-0; IP-Nav-0; IP-Nog-17; IP-Orb-10; IP-Oso-0; IP-Pal-0; IP-Pan-0; IP-Pds-1; IP-Pob-0; IP-Pro-0; IP-Pue-0; IP-Rds-0; IP-Rei-0.9510; IP-Rei-0.9574; IP-Ren-6; IP-Rev-0; IP-Ria-0; IP-Sac-0; IP-San-10; IP-Scm-0; IP-Sdv-3; IP-Ses-0; IP-Sne-0; IP-Stp-0; IP-Svi-0; IP-Tam-0; IP-Tdc-0; IP-Tol-7; IP-Tor-1; IP-Trs-0; IP-Vad-0; IP-Vae-2; IP-Vav-0; IP-Vaz-0; IP-Vdm-0; IP-Vdt-0; IP-Ver-5; IP-Vid-1; IP-Vig-1; IP-Vim-0; IP-Vin-0; IP-Vis-0; IP-Voz-0; IP-Vpa-1; ISS-20; IST-29; Iasi-1; In-0; Is-0; Istisu-1; Je-0; Jea; Jl-3; Jm-0; K-oze-1; K-oze-3; KBG1-14; KBG2-13; KYC-33; Kaevlinge-1; Kal-2; Kar-1; Karag-1; Karag-2; Kas-1; Kas-2; Kastel-1; Kb-0; Kelsterbach-4; Kent; Kia-1; Kil-0; Kin-0; Kl-5; Kn-0; Kni-1; Knjas-1; Knox-18; Ko-2; Koch-1; Kolar-1; Kolar-2; Koln; Kolyv-2; Kolyv-3; Kolyv-5; Kolyv-6; Kondara; Kor-3; Koren-1; Kro-0.MPI; Kro-0.SALK; Krot-0; Kru-3; Kulturen-1; Kus2-2; Kyoto; Kz-9; LDV-18; LDV-46; LEC-25; LI-OF-065; LL-0; LP3413.41; La-0; Lag1-2; Lag1-4; Lag1-6; Lag2; Lan-1; Lan-0; Le-0; Lebja-2; Lebja-4; Leo-1; Ler-0; Ler-1.MPI; Ler-1.SALK; Lerik1-3; Leska-1; Lesno-1; Lesno-2; Lesno-4; Li-7; Li-2; Liarum; Lilloe-1; Lip-0; Liri-1; Lis-2; Lis-3; Lisse; Litva; Lm-2; Lom1-1; Lp2-2; Lp2-6; Lu-1; Lu3-30; Lu4-2; Lund; MAR-4-16; MAR2-3; MIC-31; MIL-2; MNF-Che-2; MNF-Jac-12; MNF-Pin-39; MNF-Pot-21; MNF-Pot-75; MNF-Riv-21; MOL-1; MOU2-25; Malii-1; Marce-1; Masl-1; Mc-0; Mdn-1; Melic-1; Melni-2; Mer-6; Mh-0; Mir-0; Mitterberg-1-180; Mitterberg-1-182; Mitterberg-1-183; Mitterberg-2-184; Mitterberg-2-185; Mitterberg-3-187; Mnz-0; Ms-0; Mt-0; Muh-2; Mv-0; Mz-0; N13; NC-6; NFA-10; NFA-8; NOZ-6; Naes-2; Nc-1; Nd-1; Nemrut-1; Neo-6; Nicas-1; Nie1-2; No-0; Nok-3; Nosov-1; Noveg-1; Noveg-2; Noveg-3; Np-0; Nw-0; Nyl-13; Nyl-2; Nyl-7; Nz-1; Ob-0; Obe1-15; Obh-13; Oede-2; Oemoe1-7; Oemoe2-1; Oer-1; Old-1; Olympia-2; Omn-1; Omn-5; Or-0; Orast-1; Ove-0; Oy-0.JGI; Oy-0.WTC; PHW-2; PHW-34; PLO-1; PLY-2-; PNA3; PT2.21; PYL-6; Panik-1; Panke-1; Parti-1; Paw-26; Ped-0; Per-1; Petergof; Pfn-10; Pfn-N2.2-6; Pi-0; Pigna-1; Pla-0; Pna-10; Pna-17; Po-0; Pog-0; Pra-6; Pro-0; Pt-0; Pu2-23; Pu2-7; Pu2-8; Puk-2; QUI-8; Qar-8a; Qui-0; RAD-21; RMX3.22; RRS-7; RRS-10; RUM-20; Ra-0; Ragl-1; Rak-2; Rakit-1; Rakit-3; Rd-0.MPI; Rd-0.SALK; Ren-1;

Ren-11; Rennes-1; Rev-1; Rev-2; Rhen-1; Ri-0; Rld-1; Rmx-A02; Rmx-A180; Roed-17-319; Rome-1; Rou-0; Rsch-4; Ru-2; Ru-N2; Ru4-16; Rubeznhoe-1; Rue3-1-31; SAUL-24; SLSP-31; SLSP-35; Sakata; San-2; Sanna-2; Sap-0; Sarno-1; Schip-1; Schl-7; Se-0; Seattle-0; Sei-0; Set-1; Sever-1; Sf-2; Sf-1; Sg-1; Sha.JGI; Sha.MPI; Si-0; Sim-1; Slavi-2; Smolj-1; Sorbo; Sp-0; Sparta-1; Spr1-2; Spr1-6; Spro-1; Spro-2; Spro-3; Sq-1; Sq-8; Sr3; Sr5; St-0; Star-8; Stara-1; Staro-1; Ste-0; Ste-2; Ste-3; Ste-4; Stiav-1; Stilo-1; Stu1-1; Stw-0; Su-0; Sus-1; T1000; T1020; T1070; T1080; T1090; T1110; T1130; T1160; T460; T470; T480; T530; T540; T550; T570; T710; T720; T740; T780; T790; T800; T840; T850; T860; T880; T900; T930; T960; T980; T990; TAA-04; TAA-14; TAA-18; TAAD-01; TAAD-03; TAAD-04; TAAD-05; TAAD-06; TAAL-03; TAAL-07; TBO-01; TDr-1; TDr-13; TDr-16; TDr-17; TDr-2; TDr-7; TDr-8; TDr-9; TEDEN-02; TEDEN-03; TFAE-06; TFAE-07; TFAE-08; TOM-04; TOM-06; TOM-07; TOU-A1-88; TOU-A1-89; TRAE-01; TRE-1; TV-10; TV-22; TV-30; TV-38; TV-7; Ta-0; Tamm-2.GMI; Tamm-2.SALK; Tamm-27; Teano-1; Teiu-2; Tgr-01; Tha-1; Ting-1; Tny-04; Toc-1; Tol-0; Tomegap-2; Tottarp-2; Ts-1; Ts-5; Tscha-1; Tsu-0; Tsu-1; Tu-B1-2; Tu-B2-3; Tu-KB-6; Tu-KS-7; Tu-NK-12; Tu-PK-7; Tu-WH; Tu-0; TueSB30-3; TueV13; TueWa1-2; Tuescha9; Tul-0; Tur-4; Ty-1; Ty-0; UKID107; UKID114; UKID63; UKID74; UKID96; UKNW06-003; UKNW06-403; UKNW06-481; UKSE06-118; UKSE06-252; UKSE06-325; UKSE06-362; UKSE06-432; UKSE06-470; UKSE06-500; UKSE06-533; UKSW06-179; UKSW06-207; UKSW06-226; UKSW06-285; UKSW06-302; UKSW06-333; UKSW06-360; UduI.6296; UduI.6390; UduI.6396; Uk-1; Klies-1; Ull-A-1; Ull2-3; Ull2-5; Ullapool-8; Uod-2; Uod-1; Uod-7; Utrecht; VED-10; Vaar-1; Vaar2-1; Vaestervik; Van-0; Vash-1; Ven-1; Vie-0; Vimmerby; Vind-1; Vinsloev; WAR; WAV-8; Wa-1; WalhaesB4; Wc-1; Wei-0; Westkar-4; Wil-2; Wil-1; Wl-0; Ws-0; Wt-5; Wu-0; Xan-1; Yeg-1; Yeg-2; Yeg-4; Yeg-5; Yeg-7; Yeg-8; Yo-0; Yst-1; Zagub-1; Zal-1; Zdarec3; ZdrI.6424; ZdrI.6434; ZdrI.6445; Zdr-1; Zu-0; Zu-1; Zupan-1; love-1; love-5;

APPENDIX F: Brassicaceae Sequences Used for *ACC1*/*ACC2* Alignments and Determination of K_a/K_s Ratios

This appendix lists the details of the genome sequences from members of the Brassicaceae members used for alignments and the determination of K_a/K_s ratios, and the ACC sequence for an outgroup, *Theobroma cacao*. Only Brassicaceae species whose genomes were fully sequenced were used for these analyses. Included data are species name, name of the *ACC1* and *ACC2* sequences, the website the sequences were obtained from, and relevant publications. Adapted from Parker et al. (2014).

Species	Sequences Analyzed	Relevant Website/Citations
Arabidopsis thaliana	ACC1 (NP_174849; At1g36160) ACC2 (NP_174850; At1g36180)	NCBI (www.ncbi.nlm.nih.gov/); TAIR 10 (www.arabidopsis.org)
Brassica rapa Chiifu-401	ACC1 (Bra036771) ACC2 (Bra018702)	Phytozome v9 (www.phytozome.net/); Brassica Database (http://brassicadb.org/brad/index.php)
Arabidopsis lyrata	ACC1 (922767) ACC2 (473714)	Phytozome v10 (http://phytozome.jgi.doe.gov/pz/portal.html); Hu, et al. (2011)
Capsella rubella	ACC1 (Carubv10011872m) ACC2 (Carubv10008063m)	Phytozome v10 (http://phytozome.jgi.doe.gov/pz/portal.html); Slotte, et al. (2013)
Leavenworthia alabamica	Obtained from Genome Sequence	CoGe (https://genomevolution.org/CoGe/); Lyons, et al. (2008)
Sisymbrium irio	Obtained from Genome Sequence	CoGe (https://genomevolution.org/CoGe/); Lyons, et al. (2008)
Boechera stricta	ACC1 (Bostr.20910s0015.1)	Phytozome v10 (http://phytozome.jgi.doe.gov/pz/portal.html); *Boechera stricta* v1.2, DOE-JGI
Aethionema arabicum	Obtained from Genome Sequence	CoGe (https://genomevolution.org/CoGe/); Lyons, et al. (2008)
Brassica rapa FPsc	ACC1 (Brara.H00605.1) ACC2 (Brara.F00455.1)	Phytozome v10 (http://phytozome.jgi.doe.gov/pz/portal.html); *Brassica rapa* FPsc v1.3, DOE-JGI
Eutrema parvulum	ACC1 (Tp1g30640) ACC2 (Tp1g30680)	CoGe (https://genomevolution.org/CoGe/); Lyons, et al. (2008)
Theobroma cacao	ACC (Thecc1EG034957t1)	Phytozome v10 (http://phytozome.jgi.doe.gov/pz/portal.html); Motamayor, et al. (2013)

APPENDIX G: Eukaryotic ACCase Sequences Used for the Original Multi-Kingdom Alignment of 20 sequences

This appendix lists the details of the eukaryotic protein sequences used for the original multi-kingdom alignment. Included data are species name, name of the *ACC1* and *ACC2* sequences, the website the sequences were obtained from, and relevant publications. Adapted from Parker et al. (2014).

Footnotes for the following table are described below:

[a] The genomic sequence for *ACC1*, obtained from Phytozome v9, was used to add conserved amino acids missing from the predicted protein sequence.

[b] The genomic sequence for cytosolic ACC, obtained from NCBI, was used to add conserved amino acids missing from the predicted protein sequence.

[c] Isoform A was chosen because all of the other isoforms are contained within.

[d] The genomic sequence for *HFA1*, obtained from Saccharomyces Genome Database, was used to add conserved, N-terminal amino acids missing from the predicted protein sequence.

Species	Sequences Analyzed	Relevant Website/Citations
Arabidopsis thaliana	ACC1 (NP_174849; At1g36160)	NCBI (www.ncbi.nlm.nih.gov/);
	ACC2 (NP_174850; At1g36180)	TAIR 10 (www.arabidopsis.org)
Brassica rapa Chiifu-401	ACC1 (Bra036771) [a]	Phytozome v9 (www.phytozome.net/);
	ACC2 (Bra018702)	Brassica Database (http://brassicadb.org/brad/index.php); Cheng et al. (2011)
Medicago truncatula	ACC (XP_03638794.1)	NCBI (www.ncbi.nlm.nih.gov/)
Triticum aestivum	Cytosolic ACC (ACD46686.1)	NCBI (www.ncbi.nlm.nih.gov/)
	Plastid ACC (ACD46683.1)	
Zea mays	Cytosolic ACC; hypothetical (AFW68888.1) [b]	NCBI (www.ncbi.nlm.nih.gov/)
	Plastid ACC (AAA80214.1)	
Homo sapiens	ACC1 (Isoform 1) (NP_942131.1)	NCBI (www.ncbi.nlm.nih.gov/)
	ACC2 (Precursor) (NP_001084.3)	
Mus musculus	ACC1 (NP_579938.2)	NCBI (www.ncbi.nlm.nih.gov/)
	ACC Beta Precursor (NP_598665.2)	
Danio rerio	ACC Alpha (NP_001258237.1)	NCBI (www.ncbi.nlm.nih.gov/)
	ACC2 Isoform X1 (XP_005165251.1)	
Drosophila melanogaster	ACC Isoform A (AAF59155.2) [c]	NCBI (www.ncbi.nlm.nih.gov/)
Saccharomyces cerevisiae	ACC1 (NP_014413.1)	NCBI (www.ncbi.nlm.nih.gov/)
	ACC HFA1 (NP_013934.1) [d]	
Schizosaccharomyces pombe	Acetyl CoA/Biotin Carboxylase (NP_593271.1)	NCBI (www.ncbi.nlm.nih.gov/)
Neurospora crassa	ACC (XP_963017.1)	NCBI (www.ncbi.nlm.nih.gov/)

APPENDIX H: UniProt and NCBI Reference IDs for Eukaryotic Sequences used in Multi-Kingdom and Plant Alignments

This appendix lists the details of the eukaryotic ACCase protein sequences obtained from the Pfam database (http://pfam.xfam.org/family/PF08326) based the presence of the central domain, and BLAST searches (https://blast.ncbi.nlm.nih.gov/Blast.cgi) with both ACC1 and ACC2 Arabidopsis protein sequences. Included data are the UniProt identification numbers, the name of the NCBI reference sequence, and the species name.

UniProt ID	NCBI Reference Sequence	Species
F4WVP2_ACREC	EGI61712.1	*Acromyrmex echinatior*
J9JZ55_ACYPI	XP_003245354	*Acyrthosiphon pisum*
B2ZGJ3_AEGTA	ACD46664.1	*Aegilops tauschii*
B2ZGK8_AEGTA	ACD46679.1	*Aegilops tauschii*
M8BWJ4_AEGTA	EMT29390.1	*Aegilops tauschii*
K9HK96_AGABB	XP_006461998	*Agaricus bisporus var. bisporus*
K5XF75_AGABU	XP_007327289	*Agaricus bisporus var. burnettii*
A0A0D4WUQ0_AGRST	AJV90958.1	*Agrostis stolonifera*
G1M4G3_AILME		*Ailuropoda melanoleuca*
G1M4I7_AILME		*Ailuropoda melanoleuca*
G1MDP5_AILME		*Ailuropoda melanoleuca*
C0NM24_AJECG	EEH07675.1	*Ajellomyces capsulatus*
C6H4H2_AJECH	EER45832.1	*Ajellomyces capsulatus*
F0U8S6_AJEC8	EGC41773.1	*Ajellomyces capsulatus*
T5C3B6_AJEDE	EQL35860.1	*Ajellomyces dermatitidis*
F2TFC1_AJEDA	EGE81934.2	*Ajellomyces dermatitidis*
C5G6W5_AJEDR	EEQ83334.1	*Ajellomyces dermatitidis*
C5JZT9_AJEDS	XP_002621459	*Ajellomyces dermatitidis*
A0A024FVX6_9STRA	CCI11186.1	*Albugo candida*
F0WQ05_9STRA	CCA23409.1	*Albugo laibachii*
H9BT72_9POAL	AFD53915.1	*Alopecurus japonicus*
H9BT73_9POAL	AFD53916.1	*Alopecurus japonicus*
B5QSK0_ALOMY	CAL63610.1	*Alopecurus myosuroides*
B5QSK1_ALOMY	CAL63611.1	*Alopecurus myosuroides*
Q8LRK2_ALOMY	CAC84161.1	*Alopecurus myosuroides*
W1NEP0_AMBTC	ERM94222.1	*Amborella trichopoda*
B6RC94_AMYRO	ABQ28729.1	*Amylomyces rouxii*
U3J6N8_ANAPL		*Anas platyrhynchos*
U3IH42_ANAPL		*Anas platyrhynchos*
G1KIU6_ANOCA		*Anolis carolinensis*
W5JD41_ANODA	ETN60739.1	*Anopheles darlingi*
Q7PQ11_ANOGA	XP_001688518	*Anopheles gambiae*
W4H3Y5_9STRA	XP_009824345	*Aphanomyces astaci*
A0A024UP41_9STRA	XP_008864013	*Aphanomyces invadans*
V9IDH8_APICE	XP_016913943	*Apis cerana*
V9IET6_APICE	XP_016913945	*Apis cerana*
H9KKX0_APIME		*Apis mellifera*
D7KM08_ARALL	XP_002891212	*Arabidopsis lyrata*
D7KM09_ARALL	XP_002891213	*Arabidopsis lyrata*
ACC1_ARATH	NP_001185143	*Arabidopsis thaliana*

ACC2_ARATH	NP_174850.4	*Arabidopsis thaliana*
E6Y6S2_ARAHY	ACO53624.1	*Arachis hypogaea*
E6Y6S3_ARAHY	ACO53625.1	*Arachis hypogaea*
E6Y6S4_ARAHY	ACO53626.1	*Arachis hypogaea*
E6Y6S5_ARAHY	ACO53627.1	*Arachis hypogaea*
H6QXH0_ARAHY	ACZ50637.1	*Arachis hypogaea*
G1X4I9_ARTOA	XP_011119401	*Arhtrobotrys oligospora*
E4UQ09_ARTGP	XP_003174626	*Arthroderma gypseum*
C5FPQ6_ARTOC	XP_002846743	*Arthroderma otae*
R9X868_ASHAC	AGO10061.1	*Ashbya aceri*
Q75EK8_ASHGO	NP_982612	*Ashbya gossypii*
M9MV47_ASHG1	AEY94722.1	*Ashbya gossypii*
A1CST3_ASPCL	XP_001267796	*Aspergillus clavatus*
B8NBR1_ASPFN	XP_002378098	*Aspergillus flavus*
G7XHM5_ASPKW	GAA86434.1	*Aspergillus kawachii*
Q1JTV6_ASPNG	CAG38356.1	*Aspergillus niger*
G3XVD5_ASPNA	EHA25188.1	*Aspergillus niger*
A2QZ87_ASPNC	XP_001395476	*Aspergillus niger*
I8TWU2_ASPO3	EIT78693.1	*Aspergillus oryzae*
Q2TZI7_ASPOR	XP_001826411	*Aspergillus oryzae*
A0A017S3W0_9EURO	EYE91708.1	*Aspergillus ruber*
Q0C9D2_ASPTN	XP_001218324	*Aspergillus terreus*
W5KIQ3_ASTMX	XP_015460289	*Astyanax mexicanus*
W5KIQ5_ASTMX		*Astyanax mexicanus*
W4WKL3_ATTCE		*Atta cephalotes*
F0YE78_AURAN	XP_009038793	*Aureococcus anophagefferns*
F0YJA4-AURAN	XP_009040492	*Aureococcus anophagefferns*
J0CWF4_AURDE	XP_007356999	*Auricularia delicata*
A0A034V3P1_BACDO	JAC37911.1	*Bactrocera dorsalis*
A0A034V5X5_BACDO	JAC37909.1	*Bactrocera dorsalis*
A0A034V813_BACDO	JAC37910.1	*Bactrocera dorsalis*
K8EIK1_9CHLO	XP_007511713	*Bathycoccus prasinos*
F4NUF3_BATDJ	XP_006675199	*Batrachochytrium dendrobatidis*
M2MQH6_BAUCO	XP_007678580	*Baudoinia compniacensis*
J4KLB2_BEAB2	XP_008602470	*Beauveria bassiana*
T1SHS3_9POAL	AGT45917.1	*Beckmannia syzigachne*
T1SHX0_9POAL	AGT45916.1	*Beckmannia syzigachne*
T1SIA4_9POAL	AGT45914.1	*Beckmannia syzigachne*
T1SJX0_9POAL	AGT45915.1	*Beckmannia syzigachne*
	XP_010683396	*Beta vulgaris subsp. vulgaris*
W6Z4P0_COCMI	XP_007690903	*Bipolaris oryzae*

W7EJM1_COCVI	XP_014556586	*Bipolaris victoriae*
W6XYP0_COCCA	XP_007713091	*Bipolaris zericola*
N1JEK1_BLUG1	CCU81673.1	*Blumeria graminis f. sp. hordei*
ACACA _BOVIN	NP_776649	*Bos taurus*
E1BGH6_BOVIN	XP_005220033	*Bos taurus*
F1MSC3_BOVIN		*Bos taurus*
R1GBU5_BOTPV	XP_007589193	*Botryosphaeria parva*
M7TMZ7_BOTF1	EMR82519.1	*Botryotinia fuckeliana*
G2YUC1_BOTF4	CCD55219.1	*Botryotinia fuckeliana*
I1I3Q4_BRADI		*Brachypodium distachyon*
I1IWF2_BRADI	XP_003581375	*Brachypodium distachyon*
Q42617_BRANA	CAA54683.1	*Brassica napus*
Q9FEH8_BRANA	CAC19876.1	*Brassica napus*
Q9FNT7_BRANA	CAC19875.1	*Brassica napus*
	XP_013592802	*Brassica oleracea var. oleracea*
	XP_013603687	*Brassica oleracea var. oleracea*
M4DQA9_BRARP		*Brassica rapa subsp. pekinensis*
M4F6R1_BRARP		*Brassica rapa subsp. pekinensis*
V5FTR8_BYSSN	GAD93096.1	*Byssochlamys spectabilis*
G0MAW6_CAEBE	EGT40685.1	*Caenorhabditis brenneri*
A8X496_CAEBR	CAP27456.2	*Caenorhabditis briggsae*
H2L0M0_CAEEL	NP_001254027	*Caenorhabditis elegans*
Q9GZI3_CAEEL	NP_001022400	*Caenorhabditis elegans*
E3MCE8_CAERE	XP_003106220	*Caenorhabditis remanei*
F6WTV0_CALJA		*Callithrix jacchus*
F6XFU0_CALJA	JAB25616.1	*Callithrix jacchus*
U3DZ12_CALJA	JAB37452.1	*Callithrix jacchus*
	XP_010500069	*Camelina sativa*
	XP_010500071	*Camelina sativa*
E2AJI5_CAMFO	XP_011259212	*Camponotus floridanus*
Q5AAM4_CANAL	XP_718624	*Candida albicans*
C4YNG3_CANAW	EEQ43196.1	*Candida albicans*
B9WKR0_CANDC	XP_002421671	*Candida dubliniensis*
Q6FKK8_CANGA	XP_449236	*Candida glabrata*
M3IQY9_CANMX	EMG48936.1	*Candida maltose*
H8WWH3_CANO9	XP_003866237	*Candida orthopsilosis*
G8BCJ0_CANPC	CCE41856.1	*Candida parapsilosis*
G3BBN5_CANTC	XP_006688363	*Candida tenuis*
C5M4L7_CANTT	XP_002546225	*Candida tropicalis*
E2RL01_CANLF	XP_005624835	*Canis lupus familiaris*
F1PZY2_CANLF	XP_005636385	*Canis lupus familiaris*

R7UHT9_CAPTE	ELU03368.1	*Capitella teleta*
Q2HXS0_CAPHI	ABC96905.1	*Capra hircus*
W9YRL0_9EURO	XP_007719432	*Capronia coronate*
W9Y8N9_9EURO	XP_007730619	*Capronia epimyces*
E9CF10_CAPO3	XP_004344271	*Capsaspora owczarzaki*
R0GKZ5_9BRAS	XP_006303734	*Capsella rubella*
R0IAN7_9BRAS	XP_006306571	*Capsella rubella*
R0IQG7_9BRAS	XP_006306570	*Capsella rubella*
H0V3L2_CAVPO	XP_003477800	*Cavia porcellus*
H0V6W7_CAVPO	XP_013007591	*Cavia porcellus*
A0A026WBM2_CERBI	EZA53363.1	*Cerapachys biroi*
W8AF13_CERCA	JAB86820.1	*Ceratitis capitate*
W8B0Q2_CERCA	JAB86821.1	*Ceratitis capitate*
M2QP27_CERS8	EMD38813.1	*Ceriporiopsis subvermispora*
G0S3L5_CHATD	XP_006692638	*Chaetomium thermophilum*
M7BGW5_CHEMY	EMP37161.1	*Chelonian mydas*
	XP_004500605	*Cicer arietinum*
F6SZW6_CIOIN		*Ciona intestinalis*
F6T0F1_CIOIN		*Ciona intestinalis*
H2YM65_CIOSA		*Ciona savignyi*
H2YM68_CIOSA		*Ciona savignyi*
H2YM69_CIOSA		*Ciona savignyi*
H2YM70_CIOSA		*Ciona savignyi*
H2YM71_CIOSA		*Ciona savignyi*
H2YM72_CIOSA		*Ciona savignyi*
V4TCA6_9ROSI	XP_006434031	*Citrus clementina*
	XP_006472643	*Citrus sinensis*
V9DN70_9EURO	XP_008721837	*Cladophialophora carrionii*
W9X7H4_9EURO	XP_007742793	*Cladophialophora psammophila*
W9WH07_9EURO	XP_007752433	*Cladophialophora yegresii*
M1VXF1_CLAP2	CCE32912.1	*Claviceps purpurea*
C4Y676_CLAL4	XP_002616419	*Clavispora lusitaniae*
J3KHY4_COCIM	XP_001247056	*Coccidioides immitis*
C5PHV9_COCP7	XP_003066257	*Coccidioides posadasii*
E9DD80_COCPS	EFW15605.1	*Coccidioides posadasii*
I0YI54_ COCSC	XP_005642617	*Coccomyxa subellipsoidea*
N4X1Q9_COCH4	XP_014074435	*Cochliobolus heterostrophus*
M2TD32_COCH5	EMD95380.1	*Cochliobolus heterostrophus*
M2SHU2_COCSN	XP_007698182	*Cochliobolus sativus*
A0A068TY93_COFCA	CDP01191.1	*Coffea canephora*
A0A010R0B5_9PEZI	XP_007590342	*Colletotrichum fioriniae*

T0LNU3_COLGC	EQB49900.1	*Colletotrichum gloeosporioides*
L2GBP1_COLGN	XP_007275252	*Colletotrichum gloeosporioides*
E3QPV0_COLGM	XP_008096897	*Colletotrichum graminicola*
N4VTH9_COLOR	ENH87237.1	*Colletotrichum orbiculare*
R7YUQ2_CONA1	XP_007780952	*Coniosporium apollinis*
D6RNI5_COPC7	XP_002910856	*Coprinopsis cinerea*
G3JUL1_CORMM	XP_006674566	*Cordyceps militaris*
W4VRL7_9DIPT	JAB58048.1	*Corethrella appendiculata*
M5AJ86_CRIGR	NP_001278985	*Cricetulus griseus*
E6R880_CRYGW	XP_003194770	*Cryptococcus gattii*
J9VTZ1_CRYNH	XP_012050363	*Cryptococcus neoformans var. grubii*
Q55QT6_CRYNB	XP_774823	*Cryptococcus neoformans var. neoformans*
Q5KFC9_CRYNJ	XP_571316	*Cryptococcus neoformans var. neoformans*
E7CCB2_CTEID	ADT82650.1	*Ctenopharyngodon idella*
E7CCB3_CTEID	ADT82651.1	*Ctenopharyngodon idella*
F2YFF6_CTEID	ADX43925.1	*Ctenopharyngodon idella*
B0WE67_CULQU	XP_001847001	*Culex quinquefasciatus*
Q39478_9STRA	AAA81471.1	*Cyclotella cryptica*
W2RRS8_9EURO	XP_008718016	*Cyphellophora europaea*
M5FZR3_DACSP	EJU03511.1	*Dacryopinax sp.*
S8ABK3_DACHA	XP_011111700	*Dactylellina haptotyla*
F1QH12_DANRE	NP_001258237	*Danio rerio*
F1QM37_DANRE	XP_009299650	*Danio rerio*
F1QX79_DANRE		*Danio rerio*
F6P055_DANRE	XP_017211600	*Danio rerio*
E9G1C9_DAPPU	EFX86656.1	*Daphnia pulex*
Q6BX58_DEBHA	XP_457211	*Debaryomyces hansenii*
I2JTC0_DEKBR	EIF46222.1	*Dekkera bruxellensis*
N6U7G4_DENPD	ENN77580.1	*Dendroctnus ponderosae*
U4UM60_DENPD	ERL95164.1	*Dendroctnus ponderosae*
K9IW06_DESRO	JAA53347.1	*Desmodus rotundus*
ACAC_DICDI	XP_636722	*Dictyostelium discoideum*
F1A0W2_DICPU	XP_003293306	*Dictyostelium purpureum*
W7IDB6_9PEZI	EWC46965.1	*Drechslerella stenobrocha*
B3MGC4_DROAN	XP_001961005	*Drosophila ananassase*
B3N9A9_DROER	XP_001970537	*Drosophila erects*
A1Z784_DROME	NP_610342	*Drosophila melanogaster*
A8DY67_DROME	NP_001097227	*Drosophila melanogaster*
Q7JV23_DROME	NP_001097226	*Drosophila melanogaster*
B4KQ74_DROMO	XP_002005267	*Drosophila mojavensis*
B4GB22_DROPE	XP_002016234	*Drosophila persimilis*

Q290Y2_DROPS	XP_001360655	*Drosophila pseudoobscura*
B4HRH5_DROSE	XP_002032844	*Drosophila sechellia*
B4LJK2_DROVI	XP_002050377	*Drosophila virilis*
B4MPP4_DROWI	XP_002063097	*Drosophila willistoni*
B4P271_DROYA	XP_002089560	*Drosophila yakuba*
E5LBD4_ECHCG	ADR32358.1	*Echinochloa crus-galli*
E5LBD5_ECHCG	ADR32359.1	*Echinochloa crus-galli*
U6LW93_9EIME	CDJ53523.1	*Eimeria brunetti*
U6JZ14_9EIME	XP_013351861	*Eimeria mitis*
U6N237_9EIME	XP_013437840	*Eimeria necatrix*
U6L2T2_EIMTE	XP_013233659	*Eimeria tenella*
	XP_010916914	*Elaeis guineensis*
A0A023JGI2_ELEIN	AHI94840.1	*Eleusine indica*
A0A023JH13_ELEIN	AHI94839.1	*Eleusine indica*
V9SC70_ELEIN	AHC53985.1	*Eleusine indica*
V9SF96_ELEIN	AHC53984.1	*Eleusine indica*
O60033_EMEND	CAA75926.1	*Emericella nidulans*
G5EAT9_EMENI	XP_663730	*Emericella nidulans*
U1G9D6_ENDPU	XP_007805723	*Endocarpon pusillum*
F6RIW1_HORSE		*Equus caballus*
F6WNE8_HORSE		*Equus caballus*
F6WVE9_HORSE		*Equus caballus*
F6Z6T7_HORSE		*Equus caballus*
F7AZ64_HORSE		*Equus caballus*
G8JPL1_ERECY	XP_003644677	*Eremothecium cymbalariae*
	XP_012829819	*Erythranthe guttatus*
A0A059BL04_EUCGR	XP_010060156	*Eucalyptus grandis*
M7SLD9_EUTLA	XP_007793722	*Eutypa lata*
H6C3D3_EXODN	XP_009158609	*Exophiala dermatitidis*
M3WDI5_FELCA		*Felis catus*
M3W9F4_FELCA	XP_003994907	*Felis catus*
J4G2Q4_FIBRA	XP_012180021	*Fibroporia radiculosa*
U3K2H1_FICAL		*Ficedula albicollis*
U3JRR0_FICAL	XP_016158464	*Ficedula albicollis*
S8ECV4_FOMPI	EPT01079.1	*Fomitopsis pinicola*
	XP_011464572	*Fragaria vesca subsp. vesca*
	XP_004299600	*Fragaria vesca subsp. vesca*
W9KLY9_FUSOX	EWZ43759.1	*Fusarium oxysporum*
W9ISM2_FUSOX	EWY97682.1	*Fusarium oxysporum*
F9FWX4_FUSOF	EGU78586.1	*Fusarium oxysporum*
X0I2S5_FUSOX	EXL83138.1	*Fusarium oxysporum f. sp. conglutinans*

N4UIT4_FUSC1	ENH69935.1	*Fusarium oxysporum f. sp. cubense*
X0JR10_FUSOX	EXM03719.1	*Fusarium oxysporum f. sp. cubense*
W9NCN1_FUSOX	EWZ99371.1	*Fusarium oxysporum f. sp. lycopersici*
J9MHL6_FUSO4		*Fusarium oxysporum f. sp. lycopersici*
X0AJT4_FUSOX	EXK41112.1	*Fusarium oxysporum f. sp. melonis*
W9Q3A7_FUSOX	EXA49032.1	*Fusarium oxysporum f. sp. pisi*
X0GBR9_FUSOX	EXL61087.1	*Fusarium oxysporum f. sp. radicis-lycopersici*
X0D1I0_FUSOX	EXK88537.1	*Fusarium oxysporum f. sp. raphani*
X0LU03_FUSOX	EXM24551.1	*Fusarium oxysporum f. sp. vasinfectum*
K3VV59_FUSPC	XP_009253069	*Fusarium pseudograminearum*
J3PKA7_GAGT3	XP_009230146	*Gaeumannomyces graminis var. tritici*
M2X7M8_GALSU	XP_005709050	*Galdieria sulphuraria*
ACAC_CHICK	NP_990836	*Gallus gallus*
F1NWT0_CHICK		*Gallus gallus*
F1P1B5_CHICK		*Gallus gallus*
G3P3N9_GASAC		*Gasterosteus aculeatus*
G3QAB5_GASAC		*Gasterosteus aculeatus*
A0A024JJW1_GEOCN	CDO56660.1	*Geotrichum candidum*
A0A024JKG0_GEOCN	CDO56882.1	*Geotrichum candidum*
S0E131_GIBF5	CCT66398.1	*Gibberella fujikuroi*
W7MAM1_GIBM7	EWG44524.1	*Gibberella moniliformis*
A0A016PFV3_GIBZA	EYB24697.1	*Gibberella zeae*
I1RR68_GIBZE	XP_011326196	*Gibberella zeae*
S3DLR5_GLAL2	XP_008079618	*Glarea lozoyensis*
H0ENM5_GLAL7	EHK99907.1	*Glarea lozoyensis*
S7QBB3_GLOTA	XP_007865325	*Gloeophyllum trabeum*
Q39849_SOYBN	AAA81578.1	*Glycine max*
Q42793_SOYBN	AAA75528.1	*Glycine max*
I1JVH6_SOYBN		*Glycine max*
I1KA18_SOYBN	XP_003526593.1	*Glycine max*
A0A0B2SKF5_GLYSO	KHN45348	*Glycine soja*
G3RLM1_GORGO		*Gorilla gorilla gorilla*
G3S1F5_GORGO		*Gorilla gorilla gorilla*
G3SJ10_GORGO		*Gorilla gorilla gorilla*
A0A0B0MG53_GOSAR	KHF99346	*Gossypium arboreum*
A0A0B0PDU5_GOSAR	KHG23110	*Gossypium arboreum*
	XP_012467895	*Gossypium raimondii*
	XP_012446737	*Gossypium raimondii*
F0XNS8_GROCL	XP_014170191	*Grosmannia clavigera*
L1J8C0_GUITH	XP_005831577	*Guillardia theta*
L1JJM5_GUITH	XP_005835324	*Guillardia theta*

E2B9B3_HARSA	EFN87719.1	*Harpegnathos saltator*
W4JZY5_9HOMO	XP_009549397	*Heterobasidion ittegulare*
ACACA_HUMAN	NP_942131	*Homo sapiens*
ACACB_HUMAN	NP_001084	*Homo sapiens*
A2NX49_HUMAN	CAA48770.1	*Homo sapiens*
A0A024R0Y2_HUMAN	XP_005257324	*Homo sapiens*
B2ZZ90_HUMAN	XP_011523005	*Homo sapiens*
F2EIZ4_HORVD	BAK07316.1	*Hordeum vulgare var. distichum*
M0VU12_HORVD		*Hordeum vulgare var. distichum*
M0VU16_HORVD		*Hordeum vulgare var. distichum*
M0WLS8_HORVD		*Hordeum vulgare var. distichum*
M0WX42_HORVD		*Hordeum vulgare var. distichum*
M4BF67_HYAAE		*Hyaloperonospora arabidopsidis*
G9P7N2_HYPAI	XP_013939921	*Hypocrea atroviridis*
G0RT06_HYPJQ	XP_006968328	*Hypocrea jecorina*
G9MWJ5_HYPVG	XP_013955356	*Hypocrea virens*
V5HPY4_IXORI	JAB77822.1	*Ixodes ricinus*
D2CFN2_JATCU	NP_001295714	*Jatripha curcas*
H2B108_KAZAF	XP_003959443	*Kazachstania africana*
J7S2I0_KAZNA	CCK72032.1	*Kazachstania naganishii*
Q6CL34_KLULA	XP_455355	*Kluyveromyces lactis*
W0TEH6_KLUMA	BAO41760.1	*Kluyveromyces marxianus*
F2QLC7_KOMPC	CCA37159.1	*Komagataella phaffii*
C4QXW1_KOMPG	XP_002490365	*Komagataella phaffii*
W6MK75_9ASCO	CDK26741.1	*Kuraishia capsulata*
B0CUD8_LACBS	XP_001875210	*Laccaria bicolor*
C5DBX3_LACTC	XP_002551722	*Lachancea thermotolerans*
H3AU80_LATCH		*Latimeria chalumnae*
A4HJT6_LEIBR	XP_001567323	*Leishmania braziliensis*
E9BN78_LEIDB	XP_003863393	*Leishmania donovani*
A4I7A2_LEIIN	XP_001467621	*Leishmania infantum*
Q4Q5W1_LEIMA	XP_001685287	*Leishmania major*
E9B297_LEIMU	XP_003877816	*Leishmania mexicana*
W5MAD1_LEPOC		*Lepisosteus oculatus*
W5MCD6_LEPOC	XP_015222854	*Lepisosteus oculatus*
E5ACZ0_LEPMJ	XP_003845821	*Leptosphaeria maculans*
A5DT41_LODEL	XP_001528007	*Lodderomyces elongisporus*
V4AGG4_LOTGI	XP_009046184	*Lottia gigantean*
G3TBG5_LOXAF		*Loxodonta africana*
G3U853_LOXAF		*Loxodonta africana*
G3SQU6_LOXAF		*Loxodonta africana*

G3UB16_LOXAF		*Loxodonta africana*
G7PUJ1_MACFA	XP_005583989	*Macaca fascicularis*
H9F7K7_MACMU	AFE70616.1	*Macaca mulatta*
F7H9G5_MACMU		*Macaca mulatta*
F7H9H5_MACMU	NP_001253707	*Macaca mulatta*
F7H9H7_MACMU	XP_014974908	*Macaca mulatta*
F7HHF6_MACMU		*Macaca mulatta*
K2SEA6_MACPH	EKG15200.1	*Macrophomina phaseolina*
G4N2L8_MAGO7	XP_003711534	*Magnaporthe oryzae*
L7J9G6_MAGOP	ELQ64871.1	*Magnaporthe oryzae*
L7HYC3_MAGOY	ELQ35277.1	*Magnaporthe oryzae*
M4G989_MAGP6		*Magnaporthe poae*
M5EBG9_MALS4	CCV00237.1	*Malassezia sympodialis*
	XP_008374995	*Malus domestica*
K1Y7U5_MARBU	XP_007288156	*Marssonina brunnea f. sp. multigermtubi*
Q40326_MEDSA	AAB42144.1	*Medicago sativa*
G8A392_MEDTR	AES85930.1	*Medicago truncatula*
G8A394_MEDTR	XP_013460845	*Medicago truncatula*
T1GP39_MEGSC		*Megaselia scalaris*
F4RPF1_MELLP	XP_007411003	*Melampsora larici-populina*
G1N324_MELGA		*Meleagris gallopavo*
G1N7J4_MELGA		*Meleagris gallopavo*
G3US43_MELGA		*Meleagris gallopavo*
E9DRS1_METAQ	XP_007806780	*Metarhizium acridum*
E9F1D9_METRA	XP_007822277	*Metarhizium anisopliae*
A5DC00_PICGU	XP_001487428	*Meyerozyma guilliermondii*
U5H9T3_USTV1	KDE05599.1	*Microbotryum violaceum*
C1FD95_MICCC	XP_002507094	*Micromonas commoda*
C1ML75_MICPC	XP_003056126	*Micromonas pusilla*
A0A022RSR6_MIMGU	XP_012829819	*Mimulus guttatus*
G7E646_MIXOS	GAA98306.1	*Mixia osmundae*
V2XQN0_MONRO	XP_007845839	*Moniliophthora roreri*
F6T1L2_MONDO		*Monodelphis domestica*
W9RSS8_9ROSA	XP_010105882	*Morus notabilis*
S2J3C8_MUCC1	EPB82652.1	*Mucor circinelloides f. circinelloides*
S2K7G3_MUCC1	EPB91338.1	*Mucor circinelloides f. circinelloides*
ACACA_MOUSE	XP_006532016	*Mus musculus*
Q6JIZ0_MOUSE	XP_006530176	*Mus musculus*
E9Q4Z2_MOUSE	XP_006530176	*Mus musculus*
M0RJH5_MUSAM		*Musa acuminate subsp. malaccensis*
T1PCN5_MUSDO	XP_011291857	*Musca domestica*

M3XWR5_MUSPF	XP_004779160	*Mustela putorius*
M3YUC2_MUSPF	XP_004747231	*Mustela putorius*
M2Z3W2_MYCFI	XP_007925129	*Mycosphaerella fijiensis*
F9X7Y0_MYCGM	XP_003853979	*Mycosphaerella graminicola*
N1PTV3_MYCP1	EME45844.1	*Mycosphaerella pini*
S7MFR6_MYOBR	EPQ02904.1	*Myotis brandtii*
L5MBZ6_MYODS	ELK35901.1	*Myotis davidii*
G1P779_MYOLU		*Myotis lucifugus*
G1PQT0_MYOLU		*Myotis lucifugus*
D2W323_NAEGR	EFC36537.1	*Naegleria gruberi*
I2CQP5_ NANGC	AFJ69228.1	*Nannochloropsis gaditana*
K7IMF1_NASVI		*Nasonia vitripennis*
G0VEM8_NAUCC	XP_003676380	*Naumovozyma castellii*
G0WFR5_NAUDC	XP_003671869	*Naumovozyma dairenensis*
C7Z7U6_NECH7	XP_003045600	*Nectria haematococca*
	XP_010261220	*Nelumbo nucifera*
	XP_010269187	*Nelumbo nucifera*
A1DGG9_NEOFI	XP_001260373	*Neosartorya fischeri*
Q4X1V2_ASPFU	XP_755201	*Neosartorya fumigata*
B0XRR7_ASPFC	EDP54403.1	*Neosartorya fumigata*
F0V8G9_NEOCL	XP_003880045	*Neospora caninum*
Q7SBL5_NEUCR	XP_963017	*Neurospora crassa*
F8MLL9_NEUT8	XP_009851467	*Neurospora tetrasperma*
G4UQT6_NEUT9	EGZ71228.1	*Neurospora tetrasperma*
	XP_009758450	*Nicotiana sylvestris*
	XP_009799608	*Nicotiana sylvestris*
	XP_009592508	*Nicotiana tomentosiformis*
	XP_009629534	*Nicotiana tomentosiformis*
G1QQC3_NOMLE		*Nomascus leucogenys*
G1QLR1_NOMLE		*Nomascus leucogenys*
W1QF46_OGAPD	XP_013935288	*Ogataea parapolymorpha*
S3BXE1_OPHP1	EPE05929.1	*Ophiostoma piceae*
I3K792_ORENI		*Oreochromis niloticus*
I3J0L9_ORENI		*Oreochromis niloticus*
I3J0M0_ORENI		*Oreochromis niloticus*
I3K791_ORENI		*Oreochromis niloticus*
F7G2P5_ORNAN		*Ornithorhynchus anatinus*
G1ST24_RABIT		*Oryctolagus cuniculus*
G1T7I3_RABIT		*Oryctolagus cuniculus*
J3M5P3_ORYBR	XP_006654231	*Oryza brachyantha*
J3N219_ORYBR	XP_015697314	*Oryza brachyantha*

I1PU52_ORYGL		*Oryza glaberrima*
I1QTS0_ORYGL		*Oryza glaberrima*
A2Y2U1_OYRSI	EAY97401.1	*Oryza sativa subsp. indica*
ACC1_ORYSJ	XP_015614129	*Oryza sativa subsp. japonica*
ACC2_ORYSJ	XP_015639213	*Oryza sativa subsp. japonica*
H2LUD9_ORYLA		*Oryzias latipes*
H2M2B0_ORYLA		*Oryzias latipes*
A4RRC3_OSTLU	XP_001415874	*Ostreococcus lucimarinus*
Q01GA9_OSTTA	XP_003074384	*Ostreococcus tauri*
H0WWB9_OTOGA		*Otolemur garnettii*
H0X9V4_OTOGA		*Otolemur garnettii*
W5NRT6_SHEEP		*Ovis aries*
ACACA_SHEEP	NP_001009256	*Ovis aries*
W5Q4L4_SHEEP		*Ovis aries*
W5Q4L5_SHEEP		*Ovis aries*
K6ZH78_PANTR	JAA04111.1	*Pan troglodytes*
K7C855_PANTR	XP_511428	*Pan troglodytes*
H2R9M5_PANTR		*Pan troglodytes*
H2Q6U2_PANTR	XP_003313981	*Pan troglodytes*
C0SAJ7_PARBP	EEH22491.1	*Paracoccidioides brasiliensis*
C1GDJ1_PARBD	XP_010760691	*Paracoccidioides brasiliensis*
C1HD90_PARBA	XP_015701399	*Paracoccidioides lutzii*
E0VSX2_PEDHC	XP_002429216	*Pediculus humanus*
K7FB52_PELSI		*Pelodiscus sinensis*
K7FXF6_PELSI	XP_014436839	*Pelodiscus sinensis*
B6H276_PENCW	XP_002558828	*Penicillium chrysogenum*
K9F6Y2_PEND1	XP_014532605	*Penicillium digitatum*
K9FYG5_PEND2	EKV13602.1	*Penicillium digitatum*
B6Q960_PENMQ	XP_002146561	*Penicillium marneffei*
S7ZLA5_PENO1	EPS29466.1	*Penicillium oxalicum*
W6PT53_PENRF	CDM27060.1	*Penicillium roqueforti*
W3WUX9_9PEZI	XP_007838820	*Pestalotiopsis fici*
S4R8H8_PETMA		*Petromyzon marinus*
B7G7S4_PHATC	XP_002183067	*Phaeodactylum tricornutum*
B7GEB5_PHATC	XP_002185458	*Phaeodactylum tricornutum*
K5WIW4_PHACS	XP_007401381	*Phanerochaete carnosa*
V7ARE3_PHAVU	XP_007136223	*Phaseolus vulgaris*
	XP_008803739	*Phoenix dactylifera*
A9RJQ8_PHYPA	XP_001754424	*Physcomitrella patens subsp. patens*
A9T358_PHYPA	XP_001773073	*Physcomitrella patens subsp. patens*
D0NZ18_PHYIT	XP_002997355	*Phytophthora infestans*

W2XGI5_PHYPR	ETP21089.1	*Phytophthora parasitica*
W2ZPS7_PHYPR	ETP49030.1	*Phytophthora parasitica*
W2QH02_PHYPN	XP_008902712	*Phytophthora parasitica*
W2H9P1_PHYPR	ETK91206.1	*Phytophthora parasitica*
H3GWA7_PHYRM		*Phytophthora ramorum*
G5A3T5_PHYSP	XP_009534296	*Phytophthora sojae*
G8Y1P2_PICSO	XP_004194741	*Pichia sorbitophila*
G8Y4L9_PICSO	XP_004195832	*Pichia sorbitophila*
L0PGI2_PNEJ8	CCJ31347.1	*Pneumocystis jiroveci*
M7NNH4_PNEMU	XP_007874997	*Pneumocystis murina*
B2AV83_PODAN	XP_001907634	*Podospora anserina*
D3BI99_POLPA	EFA78999.1	*Polysphondylium pallidum*
	XP_011006151	*Populous euphratica*
	XP_011027682	*Populous euphratica*
B9GUK0_POPTR	XP_002302277	*Populous trichocarpa*
B9H763_POPTR	XP_002306591	*Populous trichocarpa*
U5GG96_POPTR	XP_006383487	*Populous trichocarpa*
U5GP90_POPTR	XP_006386394	*Populous trichocarpa*
H3E7I6_PRIPA		*Pristionchus pacificus*
	XP_008234004	*Prunus mume*
M5XVG9_PRUPE	XP_007221936	*Prunus persica*
L8FT33_PSED2	XP_012744795	*Pseudogymnoascus destructans*
M9MH78_PSEA3	GAC77683.1	*Pseudozyma antarctica*
W3VHK3_PSEA5	ETS60252.1	*Pseudozyma aphidis*
V5EU37_PSEBG	XP_016293854	*Pseudozyma brasiliensis*
R9P0W6_PSEHS	XP_012188366	*Pseudozyma hubeiensis*
L5JSN8_PTEAL	ELK01766.1	*Pteropus alecto*
E3KVF5_PUCGT	XP_003332669	*Puccinia graminis f. sp. tritici*
J3Q6D5_PUCT1		*Puccinia triticina*
E3RX86_PYRTT	XP_003302241	*Pyrenophora teres f. teres*
B2VTF1_PYRTR	XP_001932248	*Pyrenophora tritici-repentis*
U4L404_PYROM	CCX10916.1	*Pyronema omphalodes*
D3ZBE2_RAT		*Rattus norvegicus*
ACACA_RAT	NP_071529	*Rattus norvegicus*
Q1HEC0_RAT	ABF48724.1	*Rattus norvegicus*
O70151_RAT	NP_446374	*Rattus norvegicus*
E9PSQ0_RAT		*Rattus norvegicus*
L7MI62_9ACAR	JAA63507.1	*Rhipicephalus pulchellus*
U9T243_RHIID	ESA02254.1	*Rhizophagus irregularis*
I1BVP2_RHIO9	EIE80272.1	*Rhizopus delemar*
T1ICG7_RHOPR		*Rhodnius prolixus*

M7XLR4_RHOT1	XP_016272252	*Rhodosporidium toruloides*
B9RJG2_RICCO	XP_002513881	*Ricinus communis*
J8PX05_SACAR	EJS41898.1	*Saccharomyces arboricola*
W7RER6_YEASX	EWH16255.1	*Saccharomyces cerevisiae*
A0A024XHX2_YEASX	EWG88892.1	*Saccharomyces cerevisiae*
W7PYC7_YEASX	EWG84123.1	*Saccharomyces cerevisiae*
A0A024XX80_YEASX	EWG93628.1	*Saccharomyces cerevisiae*
A0A024Y0W9_YEASX	EWG94141.1	*Saccharomyces cerevisiae*
ACAC_YEAST	NP_014413	*Saccharomyces cerevisiae*
HFA1_YEAST	NP_013934	*Saccharomyces cerevisiae*
B5VPX5_YEAS6	EDZ70018.1	*Saccharomyces cerevisiae*
B5VR47_YEAS6	EDZ69596.1	*Saccharomyces cerevisiae*
E7KHI7_YEASA	EGA73179.1	*Saccharomyces cerevisiae*
N1P4Q3_YEASC	EIW08105.1	*Saccharomyces cerevisiae*
N1NXK1_YEASC	EIW08474.1	*Saccharomyces cerevisiae*
C7GLN9_YEAS2	EEU08272.1	*Saccharomyces cerevisiae*
HFA1_YEAS2	EEU06674.1	*Saccharomyces cerevisiae*
G2WL73_YEASK	GAA26109.1	*Saccharomyces cerevisiae*
G2WKR3_YEASK	GAA25656.1	*Saccharomyces cerevisiae*
C8ZFP3_YEAS8	CAY82209.1	*Saccharomyces cerevisiae*
HFA1_YEAS8	CAY82038.1	*Saccharomyces cerevisiae*
HFA1_YEAS1	EDV11698.1	*Saccharomyces cerevisiae*
B3LPM6_YEAS1	EDV12250.1	*Saccharomyces cerevisiae*
A6ZS90_YEAS7	EDN62822.1	*Saccharomyces cerevisiae*
HFA1_YEAS7	EDN64143.1	*Saccharomyces cerevisiae*
E7QK49_YEASZ	EGA84966.1	*Saccharomyces cerevisiae*
A0A023ZGW9_YEASX	AHY77104.1	*Saccharomyces cerevisiae*
A0A023ZF06_YEASX	AHY76661.1	*Saccharomyces cerevisiae*
H0H0H7_SACCK	EHN00464.1	*Saccharomyces cerevisiae* x *S. kudriavzevii*
H0GLB5_SACCK	EHN05444.1	*Saccharomyces cerevisiae* x *S. kudriavzevii*
F2U425_SALR5	XP_004996552	*Salpingoeca rosetta*
G3W9V7_SARHA		*Sarcophilus harrisii*
A3GH39_PICST	XP_001386775	*Scheffersomyces stipitis*
G4VJ84_SCHMA	CCD79485.1	*Schistosoma mansoni*
D8Q0Q3_SCHCM	XP_003032695	*Schizophyllum commune*
S9XAW9_SCHCR	XP_013024217	*Schizosaccharomyces cryophilus*
B6K3W9_SCHJY	XP_002174469	*Schizosaccharomyces japonicus*
S9R9B4_SCHOY	XP_013016185	*Schizosaccharomyces octosporus*
ACAC_SCHPO	NP_593271	*Schizosaccharomyces pombe*
W9C0L4_9HELO	ESZ90327.1	*Sclerotinia borealis*
A7EM01_SCLS1	XP_001592109	*Sclerotinia sclerotiorum*

D8SW33_SELML	XP_002987586	*Selaginella moellendorffii*
D8SWL6_SELML	XP_002987673	*Selaginella moellendorffii*
F8Q8W1_SERL3	EGN95016.1	*Serpula lacrymans var. lacrymans*
F8P7V4_SERL9	XP_007322478	*Serpula lacrymans var. lacrymans*
	XP_011083399	*Sesamum indicum*
Q84TQ5_SETIT	AAO62903.1	*Setaria italica*
Q84TQ6_SETIT	NP_001267734	*Setaria italica*
Q947M6_SETIT	AAL02056.1	*Setaria italica*
K3Y4M1_SETIT	XP_012702632	*Setaria italica*
K4A4N3_SETIT	XP_004983244	*Setaria italica*
B5QSJ9_SETVI	CAL63609.1.	*Setaria viridis*
R0IHL8_SETT2	XP_008027230	*Setosphaeria turcica*
	XP_004252541	*Solanum lycopersicum*
M1AG30_SOLTU	XP_006360278	*Solanum tuberosum*
F7WC81_SORMK	XP_003344021	*Sordaria macrospora*
C5YD68_SORBI	XP_002446178	*Sorghum bicolor*
C5YP96_SORBI	XP_002442242	*Sorghum bicolor*
G3AJ35_SPAPN	XP_007374131	*Spathaspora passalidarum*
I3M0I9_SPETR		*Spermophilus tridecemlineatus*
I3M5C3_SPETR		*Spermophilus tridecemlineatus*
M3D4W7_SPHMS	XP_016761371	*Sphaerulina musiva*
E6ZP99_SPORE	CBQ69056.1	*Sporisorium reilianum*
U7PZF9_SPOS1	ERT01014.1	*Sporothrix schenckii*
A5Z221_PIG	ABQ85554.1	*Sus scrofa*
B0LJD0_PIG	NP_001107741	*Sus scrofa*
D2D0D8_PIG	ACM42414.1	*Sus scrofa*
C9W109_PIG	ACL80208.1	*Sus scrofa*
F1RGB5_PIG	NP_001193328	*Sus scrofa*
F1S1B5_PIG		*Sus scrofa*
H0ZA42_TAEGU		*Taeniopygia guttata*
H0ZD19_TAEGU		*Taeniopygia guttata*
H2TKQ8_TAKRU		*Takifugu runripes*
H2TKQ9_TAKRU		*Takifugu runripes*
H2TKR0_TAKRU		*Takifugu runripes*
H2URL5_TAKRU		*Takifugu runripes*
H2URL6_TAKRU		*Takifugu runripes*
H2URL7_TAKRU		*Takifugu runripes*
H2URL8_TAKRU		*Takifugu runripes*
B8M2J0_TALSN	XP_002478864	*Talaromyces stipitatus*
R4XAK5_TAPDE	CCG82864.2	*Taphrina deformans*
	XP_010538957	*Tarenaya hassleriana*

M9QTR5_TETUR	AGI59311.1	*Tetranychus urticae*
M9QV47_TETUR	NP_001310078	*Tetranychus urticae*
V9LL82_TETUR	AFQ61042.1	*Tetranychus urticae*
T1KU54_TETUR	NP_001310078	*Tetranychus urticae*
Q4RSU6_TETNG	CAG08536.1	*Tetraodon nigroviridis*
H3C3C2_TETNG		*Tetraodon nigroviridis*
H3C4M0_TETNG		*Tetraodon nigroviridis*
H3CZJ8_TETNG		*Tetraodon nigroviridis*
H3DEN7_TETNG		*Tetraodon nigroviridis*
H3CZJ9_TETNG		*Tetraodon nigroviridis*
I2H6X2_TETBL	XP_004181643	*Tetrapisispora blattae*
G8BT37_TETPH	XP_003685442	*Tetrapisispora phaffii*
G8BWH2_TETPH	XP_003686857	*Tetrapisispora phaffii*
B5YMF5_THAPS	XP_002296083	*Thalassiosira pseudonana*
B8BVD1_THAPS	XP_002287470	*Thalassiosira pseudonana*
L8X4Y8_THACA	ELU43694.1	*Thanatephorus cucumeris*
M5BQ58_THACB	CCO29256.1	*Thanatephorus cucumeris*
A0A061FFG4_THECC	EOY16075.1	*Theobroma cacao*
G2Q771_THIHA	XP_003660894	*Thielavia heterothallica*
G2R9M8_THITE	XP_003655052	*Thielavia terrestris*
R8BR19_TOGMI	XP_007913516	*Togninia minima*
G8ZN64_TORDC	XP_003679269	*Torulaspora delbrueckii*
D2A5X8_TRICA	XP_008194742	*Tribolium castaneum*
E5SWR6_TRISP	XP_003369594	*Trichinella spiralis*
A0A024S9D4_HYPJR	ETS01919.1	*Trichoderma reesei*
F2PM07_TRIEC	EGE02925.1	*Trichophyton equinum*
A0A022USC1_9EURO	EZF36413.1	*Trichophyton interdigitale*
A0A059J926_9EURO	KDB23987.1	*Trichophyton interdigitale*
A0A022V308_TRIRU	EZF40380.1	*Trichophyton rubrum*
A0A023AA68_TRIRU	EZG15189.1	*Trichophyton rubrum*
A0A022VX96_TRIRU	EZF50887.1	*Trichophyton rubrum*
A0A022WU48_TRIRU	EZF61603.1	*Trichophyton rubrum*
A0A028JIW5_TRIRU	EZG04649.1	*Trichophyton rubrum*
A0A059JWR6_TRIRU	KDB32108.1	*Trichophyton rubrum*
A0A022YJK6_TRIRU	EZF83025.1	*Trichophyton rubrum*
A0A022ZFA7_TRIRU	EZF93571.1	*Trichophyton rubrum*
A0A022THY9_TRIRU	EZF16244.1	*Trichophyton rubrum*
F2SK61_TRIRC	XP_003236369	*Trichophyton rubrum*
A0A022XNS8_TRISD	EZF72144.1	*Trichophyton soudanense*
F2RQG9_TRIT1	EGD93568.1	*Trichophyton tonsurans*
J6EMY4_TRIAS	XP_014176517	*Trichosporon asahii var. asahii*

B2ZGJ6_WHEAT	ACD46667.1	*Triticum aestivum*
B2ZGK3_WHEAT	ACD46674.1	*Triticum aestivum*
B2ZGL2_WHEAT	ACD46683.1	*Triticum aestivum*
B2ZGL3_WHEAT	ACD46684.1	*Triticum aestivum*
B2ZGL4_WHEAT	ACD46685.1	*Triticum aestivum*
B2ZGL5_WHEAT	ACD46686.1	*Triticum aestivum*
Q41511_WHEAT	AAA19970.1	*Triticum aestivum*
Q41525_WHEAT	AAC49275.1	*Triticum aestivum*
O48959_WHEAT	AAC39330.1	*Triticum aestivum*
B2ZGK1_TRITD	ACD46672.1	*Triticum turgidum subsp. durum*
B2ZGL0_TRITD	ACD46681.1	*Triticum turgidum subsp. durum*
B2ZGL1_TRITD	ACD46682.1	*Triticum turgidum subsp. durum*
B2ZGJ9_TRIUA	ACD46670.1	*Triticum urartu*
B2ZGK6_TRIUA	ACD46677.1	*Triticum urartu*
M7ZJ50_TRIUA	EMS59656.1	*Triticum urartu*
Q57YR7_TRYB2	XP_847540	*Trypanosoma brucei*
C9ZWK0_TRYB9	XP_011776065	*Trypanosoma brucei*
F9WIJ0_TRYCI	CCD17138.1	*Trypanosoma congolense*
G0USX5_TRYCI	CCC92488.1	*Trypanosoma congolense*
V5BGP0_TRYCR	ESS63598.1	*Trypanosoma cruzi*
K2NT98_TRYCR	EKF38261.1	*Trypanosoma cruzi*
K4E6Y9_TRYCR	EKG06140.1	*Trypanosoma cruzi*
G0TZM7_TRYVY	CCC50055.1	*Trypanosoma vivax*
C4JEF0_UNCRE	XP_002541275	*Uncinocarpus reesii*
I2FMZ2_USTH4	CCF48285.1	*Ustilago hordei*
Q12721_USTMD	CAA86983.1	*Ustilago maydis*
Q4P5I4_USTMA	XP_760776	*Ustilago maydis*
A7TDL1_VANPO	XP_001647339	*Vanderwaltozyma polyspora*
G2X095_VERDV	XP_009652053	*Verticillium dahliae*
A0A0L9UFI5_PHAAN	KOM41670.1	*Vigna angularis*
A5AIC1_VITVI	CAN64563.1	*Vitis vinifera*
F6H0V3_VITVI	CCB45550.1	*Vitis vinifera*
R9AGJ6_WALI9	XP_009268046	*Wallemia ichthyophaga*
I4YJ49_WALSC	XP_006955825	*Wallemia sebi*
K0KVW0_WICCF	CCH45634.1	*Wickerhamomyces ciferrii*
B5DEA0_XENTR	NP_001131086	*Xenopus tropicalis*
F6URD7_XENTR		*Xenopus tropicalis*
F6URI4_XENTR		*Xenopus tropicalis*
F6URY7_XENTR		*Xenopus tropicalis*
M3ZEB6_XIPMA		*Xiphophorus maculatus*
M4A0A1_XIPMA		*Xiphophorus maculatus*

Q6CC91_YARLI	XP_501721	*Yarrowia lipolytica*
K7TS88_MAIZE	XP_008663055	*Zea mays*
Q41743_MAIZE	NP_001105373	*Zea mays*
Q7XYR3_MAIZE	AAP78897.1	*Zea mays*
Q7XYR4_MAIZE	AAP78896.1	*Zea mays*
A0A0K9PL54_ZOSMR	KMZ68975.1	*Zostera marina*
W0VNJ2_ZYGBA	CDH11002.1	*Zygosaccharomyces bailii*
W0W6G6_ZYGBA	CDH17251.1	*Zygosaccharomyces bailii*
S6EL55_ZYGB2	CDF91322.1	*Zygosaccharomyces bailii*
B2G4R2_ZYGRO	CAQ43571.1	*Zygosaccharomyces rouxii*
C5DVR9_ZYGRC	XP_002496821	*Zygosaccharomyces rouxii*

APPENDIX I: Accession Consensus ACC2 Protein Sequence Reflecting Genetic Variation and Level of Conservation

This appendix shows the accession consensus protein sequences for ACC1 and ACC2 along with the variation found among all 857 accessions, conservation of each residue, and the current classification of each variant. Included data are the position number of amino acid consensus sequence for ACC2 and ACC1; variants found in ACC2 and ACC1; the number of accessions where each variant is found; the percent conservation for each ACC2 residue based on three alignments: (1) the original multi-kingdom alignment of 20 eukaryotic sequences, (2) the alignment of 139 plant sequences, and (3) the multi-kingdom alignment of 667 eukaryotic sequences; the protein domains, and the classification of each variant based on known information about it.

Footnotes for the title row of the following table are described below:

[a] Letter at a number indicates that a different amino acid than the accession consensus is the most common among the alignment. For example, "G at 10.8" means "G" (glycine) is the most common amino acid in the plant alignment with 10.8% conservation. Red numbers, ≥ 99% conserved; Purple numbers, ≥ 95% and < 99%; Blue numbers, ≥ 90% and < 95%; Green numbers, ≥ 80% and < 90%; Black numbers, < 80%.

[b] TP, transit peptide domain; BC, biotin carboxylase; BCCP, biotin carboxyl carrier protein; CEN, central domain; CT-β, carboxyltransferase-beta subunit; CT-α, carboxyltransferase-alpha subunit.

[c] D, deleterious to protein function; LD, likely deleterious; PD, possibly deleterious; VUS, variant of unknown significance; LND, likely not deleterious; ND, not deleterious.

ACC2 Protein Sequence				ACC1 Protein Sequence				% Conservation (based on ACC2) [a]			Domain [b]	Variant Classification [c]
Position	Accession Consensus	Substitution	Num. of Accessions	Position	Accession Consensus	Substitution	Num. of Accessions	Original (20)	Plant (139)	MUSCLE (667)		
1	M			-	-	-	-	10	G at 10.8	D at 10.6	TP	-
2	E			-	-	-	-	10	V at 25.2	L at 13.9	TP	-
3	M			-	-	-	-	15	S at 20.1	L at 8.4	TP	-
4	R	T	2	-	-	-	-	10	D at 26.6	E at 14.2	TP	VUS
5	A			-	-	-	-	10	15.1	E at 11.1	TP	-
6	L	S	2	-	-	-	-	10	K at 10.8	< 5	TP	VUS
7	G	V	175	-	-	-	-	5	K at 12.2	< 5	TP	VUS
8	S			-	-	-	-	10	H at 17.3	R at 15.7	TP	-
9	S			-	-	-	-	5	N at 9.4	17.5	TP	-
10	C			-	-	-	-	10	Q at 21.6	L at 13.9	TP	-
11	S			-	-	-	-	5	29.5	< 5	TP	-
12	T			-	-	-	-	5	I at 16.6	< 5	TP	-
13	G			-	-	-	-	10	R at 26.6	9.6	TP	-
14	N			-	-	-	-	10	Q at 24.5	11.2	TP	-
15	G			-	-	-	-	10	33.8	14.7	TP	-
16	G			-	-	-	-	10	< 5	< 5	TP	-
17	S			-	-	-	-	15	6.5	< 5	TP	-
18	A	T	40	-	-	-	-	5	D at 5.8	< 5	TP	VUS
19	P			-	-	-	-	5	< 5	< 5	TP	-
20	I			-	-	-	-	5	< 5	< 5	TP	-
21	T			-	-	-	-	10	< 5	< 5	TP	-
22	L			-	-	-	-	10	5.8	< 5	TP	-
23	T			-	-	-	-	15	< 5	< 5	TP	-
24	N			-	-	-	-	20	5.8	< 5	TP	-
25	I			-	-	-	-	5	< 5	< 5	TP	-
26	S			-	-	-	-	10	7.9	< 5	TP	-
27	P			-	-	-	-	10	6.5	< 5	TP	-
28	W			-	-	-	-	10	6.5	< 5	TP	-
29	I			-	-	-	-	15	6.5	< 5	TP	-
30	T			-	-	-	-	15	7.2	< 5	TP	-
31	T			-	-	-	-	5	< 5	< 5	TP	-
32	V			-	-	-	-	10	< 5	< 5	TP	-
33	F			-	-	-	-	5	L at 26.6	< 5	TP	-
34	P			-	-	-	-	10	A at 27.3	< 5	TP	-
35	S			-	-	-	-	5	G at 27.3	< 5	TP	-

36	T			-	-	-	-	10	I at 28.1	< 5	TP	-
37	V			-	-	-	-	10	I at 27.3	< 5	TP	-
38	K			-	-	-	-	15	D at 27.3	< 5	TP	-
39	L			-	-	-	-	15	33.1	< 5	TP	-
40	R			-	-	-	-	5	P at 24.5	< 5	TP	-
41	S			-	-	-	-	15	E at 21.6	< 5	TP	-
42	S			-	-	-	-	20	A at 16.6	< 5	TP	-
43	L			-	-	-	-	15	R at 15.8	< 5	TP	-
44	R			-	-	-	-	20	A at 19.4	< 5	TP	-
45	T			-	-	-	-	10	P at 19.4	< 5	TP	-
46	F			-	-	-	-	10	M at 28.1	< 5	TP	-
47	K			-	-	-	-	20	V at 30.2	< 5	TP	-
48	G			-	-	-	-	10	D at 33.1	< 5	TP	-
49	V			-	-	-	-	10	I at 36.0	< 5	TP	-
50	S			-	-	-	-	15	34.5	< 5	TP	-
51	S			-	-	-	-	10	H at 27.3	< 5	TP	-
52	R			-	-	-	-	10	G at 43.9	< 5	TP	-
53	V			-	-	-	-	10	N at 25.9	< 5	TP	-
54	R			-	-	-	-	10	E at 26.6	< 5	TP	-
55	T			-	-	-	-	5	D at 21.6	< 5	TP	-
56	F			-	-	-	-	5	P at 20.1	< 5	TP	-
57	K			-	-	-	-	10	R at 25.2	< 5	TP	-
58	G			-	-	-	-	5	20.1	< 5	TP	-
59	V	L	2	-	-	-	-	5	P at 13.0	< 5	TP	VUS
60	S			-	-	-	-	25	< 5	< 5	TP	-
61	S			-	-	-	-	10	< 5	< 5	TP	-
62	T			-	-	-	-	5	< 5	< 5	TP	-
63	R			-	-	-	-	5	< 5	< 5	TP	-
64	V			-	-	-	-	20	< 5	< 5	TP	-
65	L			-	-	-	-	10	< 5	< 5	TP	-
66	S	F	58	-	-	-	-	10	< 5	< 5	TP	VUS
67	R			-	-	-	-	5	6.5	< 5	TP	-
68	T			-	-	-	-	5	< 5	S at 16.3	TP	-
69	K			-	-	-	-	10	6.5	P at 11.1	TP	-
70	Q			-	-	-	-	15	6.5	A at 17.2	TP	-
71	Q			-	-	-	-	10	< 5	S at 15.9	TP	-
72	F			-	-	-	-	10	< 5	V at 9.0	TP	-
73	P			-	-	-	-	5	6.5	< 5	TP	-
74	L			-	-	-	-	5	< 5	< 5	TP	-
75	F			-	-	-	-	5	< 5	L at 8.7	TP	-
76	C			-	-	-	-	5	< 5	S at 9.8	TP	-

77	F			-	-	-	-	5	< 5	S at 13.6	TP	-
78	L			-	-	-	-	20	< 5	D at 11.8	TP	-
79	N			-	-	-	-	10	< 5	G at 12.4	TP	-
80	P			-	-	-	-	10	< 5	< 5	TP	-
81	D			-	-	-	-	35	6.5	N at 13.2	TP	-
82	P			-	-	-	-	15	6.5	G at 14.1	TP	-
83	I			-	-	-	-	5	< 5	L at 8.6	TP	-
84	S			-	-	-	-	30	5	Q at 8.9	TP	-
85	F			-	-	-	-	10	6.5	G at 12.6	TP	-
86	L			-	-	-	-	15	< 5	S at 9.3	TP	-
87	D	E	333	-	-	-	-	5	< 5	S at 14.2	TP	VUS
88	N			-	-	-	-	10	6.5	D at 18.1	TP	-
89	D			-	-	-	-	25	6.5	Y at 22.8	TP	-
90	V			-	-	-	-	5	< 5	A at 18.9	TP	-
91	S	C	1	-	-	-	-	20	6.5	A at 24.3	TP	VUS
92	E			-	-	-	-	10	6.5	K at 22.3	-	-
93	A			-	-	-	-	5	< 5	H at 38.1	-	-
94	E			-	-	-	-	5	< 5	M at 10.0	-	-
95	R			-	-	-	-	5	< 5	26.4	-	-
96	T			-	-	-	-	15	< 5	L at 24.7	-	-
97	V			-	-	-	-	5	< 5	S at 30.3	-	-
98	V			-	-	-	-	5	< 5	M at 29.5	-	-
99	L			-	-	-	-	15	< 5	S at 27.1	-	-
100	P			-	-	-	-	10	< 5	G at 25.2	-	-
101	D	G	82	-	-	-	-	5	< 5	L at 18.3	-	VUS
102	G			-	-	-	-	5	< 5	H at 52.9	-	-
103	S			-	-	-	-	20	< 5	F at 38.7	-	-
104	V	A	2	-	-	-	-	5	< 5	I at 39.3	-	VUS
105	N			-	-	-	-	5	< 5	K at 19.3	-	-
106	G			1	M			10	< 5	Q at 13.0	-	LND
107	A			2	A			20	< 5	G at 31.2	-	-
108	G			3	G			45	< 5	< 5	-	-
109	S			4	S			25	< 5	< 5	-	-
110	V			5	V			15	< 5	R at 20.5	-	-
111	N			6	N			45	S at 38.9	D at 14.7	-	-
112	G	V	46	7	G	R	2	60	49.6	R at 16.6	-	ND
113	Y	C	3	8	N			15	Q at 42.5	K at 13.5	-	LND
114	H			9	H			35	M at 31.7	< 5	-	-
115	S			10	S			10	N at 80.6	< 5	-	-
116	D			11	A			5	G at 75.5	< 5	-	LND
117	V			12	V			15	26.6	G at 40.2	-	-

118	V			13	G	R	1	5	H at 25.9	G at 30.4	-	LND
119	P			14	P			20	N at 36.0	N at 39.3	-	-
120	G			15	G	D	1	15	35.3	S at 20.2	-	LND
121	R			16	I			15	77.7	L at 27.1	-	LND
122	N			17	N			15	H at 43.9	R at 29.4	-	-
-	-	-	-	18	Y			-	-	-	-	LND
-	-	-	-	19	E			-	-	-	-	LND
-	-	-	-	20	T			-	-	-	-	LND
123	V			21	V			55	26.6	28.3	-	-
124	A			22	S			40	S at 86.3	S at 47.1	-	LND
125	E			23	Q			20	39.6	S at 28.5	-	LND
126	V			24	V			50	93.5	52.8	-	-
127	N			25	D			5	D at 57.6	K at 27.7	-	LND
128	E			26	E			80	83.5	52.8	-	-
129	F	L	1	27	F			90	89.9	84.9	-	VUS
130	C			28	C			45	99.3	V at 70.6	-	-
131	K			29	K			35	27.3	22.9	-	-
132	A	V (S)	83 (82)	30	A			45	85.6	27.9	-	VUS
133	L			31	L			40	96.4	F at 27.3	-	-
134	G			32	R	G	261	85	95.7	66.1	-	ND
135	G	E	1	33	G			95	100	95.7	-	LD
136	K			34	K			30	64.8	H at 40.3	-	-
137	R			35	R			40	31.6	T at 34.9	-	-
138	P			36	P			45	96.4	V at 67.2	BC	-
139	I			37	I			100	100	96.6	BC	-
140	H			38	H			45	92.1	T at 29.2	BC	-
141	S			39	S			45	97.8	K at 44.4	BC	-
142	I			40	I			45	65.5	V at 69.3	BC	-
143	L			41	L			100	100	97.9	BC	-
144	V			42	I			30	51.8	I at 83.7	BC	LND
145	A			43	A			100	97.8	95.8	BC	-
146	T			44	N			10	N at 93.5	N at 95.5	BC	LND
147	N			45	N			100	100	97.8	BC	-
148	G			46	G			100	100	97.6	BC	-
149	M			47	M			45	98.6	I at 71.4	BC	-
150	A			48	A			100	100	95.8	BC	-
151	A			49	A			100	99.3	97.6	BC	-
152	V			50	V			90	71.9	88.8	BC	-
153	K			51	K			100	100	96.9	BC	-
154	F			52	F			45	97.8	E at 39.0	BC	-
155	I			53	I			40	M at 52.5	55.5	BC	-

156	R			54	R			100	99.3	93.6	BC	-
157	S			55	S			100	100	97.3	BC	-
158	V			56	V			45	63.3	48.4	BC	-
159	R			57	R			100	100	98.5	BC	-
160	T	A	2	58	T			30	78.4	K at 39.4	BC	VUS
161	W			59	W			100	97.1	97.8	BC	-
162	A			60	A			80	95.7	71.5	BC	-
163	Y			61	Y			80	54	87.7	BC	-
164	E			62	E			90	72.7	86.4	BC	-
165	T			63	T			65	97.1	60.7	BC	-
166	F			64	F			100	98.6	95.2	BC	-
167	G	D	32	65	G			55	100	64	BC	VUS
168	S			66	T			20	T at 48.2	N at 41.1	BC	LND
169	E			67	E			95	95	87.7	BC	-
170	K			68	K			45	91.4	R at 65.5	BC	-
171	A			69	A			80	97.1	83.4	BC	-
172	V	I	2	70	I			10	I at 89.9	I at 80.4	BC	LND
173	K			71	L			15	L at 62.6	Q at 34.2	BC	LND
174	L			72	L			45	99.3	F at 78.1	BC	-
175	V			73	V			80	71.2	60.6	BC	-
176	A			74	G			40	93.5	V at 68.2	BC	LND
177	M			75	M			100	100	99.4	BC	-
178	A			76	A	T	25	65	100	69.6	BC	ND
179	T			77	T			100	100	98.4	BC	-
180	P			78	P			100	98.6	97.9	BC	-
181	E			79	E			90	99.3	95.7	BC	-
182	D			80	D			100	100	98.7	BC	-
183	M			81	M			35	77.7	L at 79.8	BC	-
184	R			82	R			45	92.1	K at 39.1	BC	-
185	I			83	I			45	98.6	A at 70.3	BC	-
186	N	I	1	84	N			100	99.3	98.7	BC	VUS
187	A			85	A			95	99.3	95.2	BC	-
188	E	D	1	86	E			90	99.3	78	BC	D
189	H			87	H			45	98.6	Y at 73.6	BC	-
190	I			88	I			100	98.6	96.4	BC	-
191	R			89	R			65	99.3	67.5	BC	-
192	I			90	I			45	92.1	M at 73.9	BC	-
193	A			91	A	V	1	100	99.3	98.4	BC	LND
194	D			92	D			100	98.6	95.7	BC	-
195	Q			93	Q			65	98.6	55.5	BC	-
196	F	L	1	94	F			45	97.1	Y at 66.7	BC	VUS

197	V			95	V			80	81.3	87	BC	-
198	E			96	E			55	96.4	63.3	BC	-
199	V			97	V			100	99.3	98.5	BC	-
200	P			98	P			100	99.3	99.6	BC	-
201	G			99	G			100	99.3	98.7	BC	-
202	G			100	G			100	99.3	99.6	BC	-
203	T			101	T			65	98.6	61.8	BC	-
204	N			102	N			100	99.3	98.5	BC	-
205	N			103	N			100	99.3	93.7	BC	-
206	N			104	N			95	95	88.3	BC	-
207	N			105	N			100	96.4	99.1	BC	-
208	Y			106	Y			100	99.3	98.4	BC	-
209	A			107	A			100	99.3	98.4	BC	-
210	N			108	N			100	99.3	99.6	BC	-
211	V			109	V			95	99.3	97.6	BC	-
212	Q			110	Q			40	95	E at 52.2	BC	-
213	L			111	L			95	97.1	93.6	BC	-
214	I			112	I			100	99.3	98.5	BC	-
215	V			113	V			70	88.5	74.8	BC	-
216	E			114	E			40	97.8	D at 66.9	BC	-
217	M			115	M			25	48.9	I at 58.3	BC	-
218	A			116	A			100	99.3	98.2	BC	-
219	E			117	E			55	87.1	62.2	BC	-
220	V			118	V			15	R at 44.6	R at 79.5	BC	-
221	T			119	T			50	85.6	33.9	BC	-
222	R			120	R			25	G at 33.8	G at 21.9	BC	-
223	V			121	V			100	99.3	95.2	BC	-
224	D			122	D			35	47.5	H at 32.7	BC	-
225	A			123	A			100	99.3	98.5	BC	-
226	V			124	V			100	96.4	96.9	BC	-
227	W			125	W			100	99.3	97.5	BC	-
228	P			126	P			45	98.6	A at 73.8	BC	-
229	G			127	G			100	99.3	99.4	BC	-
230	W			128	W			100	99.3	99.4	BC	-
231	G			129	G			100	99.3	99.4	BC	-
232	H			130	H			100	99.3	98.1	BC	-
233	A			131	A			100	99.3	98.5	BC	-
234	S			132	S			100	99.3	98.5	BC	-
235	E			133	E			100	99.3	99.4	BC	-
236	N			134	N			100	92.8	96.7	BC	-
237	P			135	P			100	98.6	98.8	BC	-

238	E			136	E			45	96.4	K at 48.4	BC	-
239	L			137	L			100	99.3	99.6	BC	-
240	P			138	P			100	99.3	98.8	BC	-
241	D			139	D			45	95.7	E at 70.0	BC	-
242	A			140	A			45	94.2	L at 29.7	BC	-
243	L			141	L			100	99.3	97.6	BC	-
244	K			142	D			20	T at 24.5	A at 41.2	BC	LND
245	E			143	A			15	A at 78.4	A at 48.9	BC	LND
246	K			144	K			40	85.6	39.1	BC	-
247	G			145	G			75	97.1	43	BC	-
248	I			146	I			80	95.7	85.5	BC	-
249	I			147	I			35	V at 55.4	V at 36.6	BC	-
250	F			148	F			100	100	99.4	BC	-
251	L			149	L			65	99.3	I at 49.5	BC	-
252	G			150	G			100	100	99.7	BC	-
253	P			151	P			100	97.1	98.7	BC	-
254	P			152	P			90	95.7	87.6	BC	-
255	A			153	A			30	71.2	G at 39.0	BC	-
256	D	A	71	154	S			5	A at 36.0	S at 38.5	BC	LND
257	S			155	S			40	77	A at 76.2	BC	-
258	M			156	M			100	100	99.1	BC	-
259	I			157	A			25	A at 52.5	R at 43.2	BC	LND
260	A			158	A			80	99.3	54.3	BC	-
261	L			159	L			100	100	99.3	BC	-
262	G			160	G			100	100	99.6	BC	-
263	D			161	D			100	100	99.3	BC	-
264	K			162	K			100	100	99.3	BC	-
265	I			163	I			80	71.2	85.9	BC	-
266	G			164	G			45	98.6	S at 42.1	BC	-
267	S			165	S			100	99.3	96.6	BC	-
268	S			166	S			60	70.5	T at 53.5	BC	-
269	L			167	L			45	99.3	I at 78.1	BC	-
270	I			168	I	R	1	45	98.6	V at 72.3	BC	LND
271	A			169	A	D	1	100	99.3	99.3	BC	VUS
272	Q			170	Q	R	1	100	100	99.3	BC	VUS
273	A			171	A			45	97.1	H at 30.9	BC	-
274	A			172	A			90	99.3	89.7	BC	-
275	D	V	1	173	D			25	G at 48.9	G at 37.8	BC	VUS
276	V	G (I)	1 (1)	174	V			65	99.3	73.3	BC	VUS
277	P			175	P			100	97.1	98.2	BC	-
278	T			176	T			80	99.3	58.8	BC	-

279	L			177	L			80	98.6	49.6	BC	-
280	P			178	P			90	76.3	84.3	BC	-
281	W			179	W			100	99.3	99.4	BC	-
282	S			180	S			100	100	97	BC	-
283	G			181	G			10	100	99.4	BC	-
284	S			182	S			75	99.3	54.8	BC	-
285	H	N	6	183	H			45	89.2	G at 66.6	BC	VUS
286	V			184	V			55	99.3	L at 28.0	BC	-
287	K			185	K			40	63.3	19	BC	-
288	I			186	I			30	63.3	V at 57.7	BC	-
289	P			187	P			45	90.7	D at 42.1	BC	-
290	P			188	P			30	43.9	W at 19.9	BC	-
291	G			189	N	S	1	10	E at 77.0	V at 33.4	BC	LND
292	R			190	S			5	S at 59.7	E at 28.0	BC	LND
293	S			191	N			15	C at 75.5	C at 15.9	BC	LND
294	L			192	L			35	79.1	I at 32.8	BC	-
295	V			193	V			20	39.6	48.9	BC	-
296	T			194	T			35	S at 42.5	38.1	BC	-
297	V	I	32	195	I			55	I at 97.8	65.1	BC	LND
298	P			196	P			85	99.3	57.3	BC	-
299	E			197	E			45	52.5	D at 48.0	BC	-
300	E			198	E			65	89.9	48.1	BC	-
301	I			199	I			30	57.6	V at 33.6	BC	-
302	Y			200	Y			95	100	90.1	BC	-
303	K			201	R	L	2	10	R at 69.1	E at 18.3	BC	ND
304	K			202	Q			50	42.5	50.1	BC	LND
305	A	V	1	203	A			50	96.4	G at 61.5	BC	VUS
306	C			204	C			90	100	84.3	BC	-
307	V			205	V			85	100	68.5	BC	-
308	Y			206	Y			25	48.2	T at 23.2	BC	-
309	T			207	T			45	97.8	S at 37.5	BC	-
310	T			208	T			45	89.2	V at 28.8	BC	-
311	E			209	E			65	84.9	63.3	BC	-
312	E			210	E			70	99.3	67.3	BC	-
313	A	V	1	211	A			50	99.3	G at 65.7	BC	VUS
314	I	V	69	212	I	T	1	20	V at 49.6	L at 73.8	BC	LND
315	A			213	A			60	96.4	E at 32.4	BC	-
316	S			214	S			50	97.8	K at 30.1	BC	-
317	C			215	C			45	99.3	A at 71.2	BC	-
318	Q			216	Q			45	95	E at 27.9	BC	-
319	V			217	V			40	76.3	E at 26.2	BC	-

320	V			218	V			55	80.6	I at 69.7	BC	-
321	G			219	G			100	100	99.9	BC	-
322	Y			220	Y			60	98.6	F at 53.5	BC	-
323	P			221	P			100	97.1	99	BC	-
324	A			222	A			45	99.3	V at 66.4	BC	-
325	M			223	M			95	100	96.7	BC	-
326	I			224	I			95	100	79.8	BC	-
327	K			225	K			100	100	100	BC	-
328	A			226	A			100	100	100	BC	-
329	S			227	S			100	100	99.1	BC	-
330	W			228	W			45	99.3	E at 77.7	BC	-
331	G			229	G			100	96.4	99.3	BC	-
332	G			230	G			100	97.1	99.3	BC	-
333	G			231	G			100	97.1	99.4	BC	D
334	G			232	G			100	100	100	BC	-
335	K			233	K			100	100	99.7	BC	-
336	G			234	G			100	100	100	BC	-
337	I			235	I			100	100	100	BC	-
338	R			236	R			100	100	100	BC	-
339	K			237	K			80	98.6	80.2	BC	-
340	V			238	V			90	100	77.7	BC	-
341	H			239	H			40	82.7	E at 31.3	BC	-
342	N			240	N			60	95	41.4	BC	-
343	D	G	30	241	D			40	92.8	E at 34.5	BC	VUS
344	D			242	D			65	95	E at 48.7	BC	-
345	E			243	E			45	92.1	D at 45.1	BC	-
346	V			244	V			45	99.3	F at 68.8	BC	-
347	R			245	R			45	84.2	P at 27.7	BC	-
348	A	G	1	246	A			55	92.8	40.6	BC	VUS
349	L			247	L			95	99.3	78	BC	-
350	F			248	F			80	97.1	58.8	BC	-
351	K			249	K			45	99.3	33.7	BC	-
352	Q	K	1	250	Q			95	98.6	67.5	BC	VUS
353	V			251	V			85	100	67.8	BC	-
354	Q			252	Q			80	99.3	52.9	BC	-
355	G	V	130	253	G			45	96.4	30.1	BC	VUS
356	E			254	E			100	100	100	BC	-
357	V			255	V			70	100	53.2	BC	-
358	P			256	P			100	100	96.1	BC	-
359	G			257	G			100	100	99.4	BC	-
360	S			258	S			100	100	98.1	BC	-

361	P			259	P			100	97.1	97.5	BC	-
362	I	T	115	260	I			90	100	89.1	BC	VUS
363	F	L	5	261	F			100	97.1	99.3	BC	PD
364	I			262	I			55	95.7	63.3	BC	-
365	M			263	M			100	100	100	BC	-
366	K			264	K			70	72.7	74.2	BC	-
367	V			265	V			40	69.8	L at 79.8	BC	-
368	A			266	A			95	100	91.3	BC	-
369	S			267	S			45	94.2	G at 37.0	BC	-
370	Q			268	Q			60	97.1	35.1	BC	-
371	S	I	1	269	S			55	97.8	A at 64.9	BC	VUS
372	R			270	R			100	100	99.4	BC	-
373	H			271	H			100	100	99.9	BC	-
374	L			272	L			100	100	97.6	BC	-
375	E			273	E			100	100	100	BC	-
376	V	A (I)	12 (1)	274	V			100	100	100	BC	PD
377	Q			275	Q			100	100	100	BC	-
378	L			276	L			65	100	72.1	BC	-
379	L			277	L			95	96.4	93.7	BC	-
380	C			278	C			45	99.3	A at 75.6	BC	-
381	D			279	D			100	100	100	BC	-
382	Q			280	K			85	69.8	78.4	BC	LND
383	Y			281	H			80	71.9	90.6	BC	LND
384	G			282	G			100	95.7	98.2	BC	-
385	N			283	N			90	100	85.2	BC	-
386	V			284	V			45	98.6	A at 41.5	BC	-
387	A			285	S			30	88.5	I at 69.6	BC	LND
388	A			286	A			45	99.3	S at 69.7	BC	-
389	L			287	L			100	100	94	BC	-
390	H			288	H			45	99.3	F at 71.2	BC	-
391	S			289	S			45	99.3	G at 75.6	BC	-
392	R			290	R			100	97.1	99.4	BC	-
393	D			291	D			100	100	100	BC	-
394	C			292	C			100	100	99.9	BC	-
395	S			293	S			100	100	99.7	BC	-
396	V	L	1	294	V			75	88.5	74.4	BC	VUS
397	Q			295	Q			100	96.4	99.3	BC	-
398	R			296	R			100	99.3	99.9	BC	-
399	R			297	R			100	99.3	99.9	BC	-
400	H			298	H			100	95.7	95.7	BC	-
401	Q			299	Q			100	100	100	BC	-

402	K			300	K			100	99.3	99.9	BC	-
403	I			301	I			100	100	99.1	BC	-
404	I	K	20	302	I			95	100	94.8	BC	LD
405	E			303	E			100	100	100	BC	-
406	E			304	E			100	100	100	BC	D
407	G			305	G			45	96.4	A at 72.4	BC	-
408	P			306	P			100	100	100	BC	-
409	I			307	I			35	70.5	V at 51.3	BC	-
410	T	N	1	308	T			85	100	79.2	BC	VUS
411	V			309	V			50	95	I at 60.1	BC	-
412	A			310	A			95	100	91.3	BC	-
413	P			311	P			50	94.2	K at 32.8	BC	-
414	Q			312	P	S	2	5	P at 33.8	P at 51.4	BC	ND
415	E			313	E			50	83.5	41.1	BC	-
416	T			314	T			65	97.8	61.3	BC	-
417	I			315	V			15	V at 82.0	F at 67.9	BC	LND
418	K			316	K			50	94.2	E at 35.2	BC	-
419	K			317	K			20	39.6	E at 27.4	BC	-
420	L			318	L			45	99.3	M at 78.9	BC	-
421	E			319	E			100	100	96.4	BC	-
422	Q			320	Q			70	98.6	38.5	BC	-
423	A			321	A			70	95	73.6	BC	-
424	A			322	A			100	98.6	99.7	BC	-
425	R	T	1	323	R			45	97.1	V at 69.9	BC	VUS
426	R			324	R			75	100	78.3	BC	-
427	L			325	L			100	100	98.7	BC	-
428	A			326	A			80	100	61.3	BC	-
429	K			327	K			85	96.4	77.8	BC	-
430	S			328	S			25	C at 38.9	L at 41.2	BC	-
431	V			329	V			100	100	97.3	BC	-
432	N			330	N			25	52.5	G at 81.4	BC	-
433	Y			331	Y			100	100	99.9	BC	-
434	V			332	V			90	84.9	89.8	BC	-
435	G			333	G			45	99.3	S at 71.4	BC	-
436	A			334	A			100	100	94.9	BC	-
437	A			335	A			45	99.3	G at 79.0	BC	-
438	T			336	T			100	100	100	BC	-
439	V			337	V			95	100	98.4	BC	-
440	E			338	E			100	100	99.9	BC	-
441	Y			339	Y			95	97.1	96.7	BC	-
442	L			340	L			100	100	97.8	BC	-

443	Y	C	1	341	Y			95	92.8	94	BC	PD
444	S			342	S			90	97.8	82	BC	-
445	M	T	190	343	M			45	99.3	H at 35.1	BC	VUS
446	D			344	D			25	E at 69.8	E at 28.8	BC	-
447	T			345	T			45	98.6	D at 55.3	BC	-
448	G			346	G			80	99.3	52.3	BC	-
449	E	D	3	347	E			35	93.5	K at 38.5	BC	VUS
450	Y			348	Y			50	95.7	F at 68.5	BC	-
451	Y			349	Y			75	95.7	76.8	BC	-
452	F			350	F			100	100	99.4	BC	-
453	L			351	L			100	100	99.9	BC	-
454	E			352	E			100	100	99.9	BC	-
455	L			353	L			100	100	99.9	BC	-
456	N			354	N			100	100	100	BC	-
457	P			355	P			100	100	100	BC	-
458	R			356	R			100	100	100	BC	-
459	L			357	L			100	100	97.9	BC	-
460	Q			358	Q			100	100	99.9	BC	-
461	V			359	V			100	99.3	99.7	BC	-
462	E			360	E			100	99.3	99.7	BC	-
463	H			361	H			100	99.3	99.7	BC	-
464	P			362	P			100	99.3	99.6	BC	-
465	V			363	V			45	98.6	T at 41.4	BC	-
466	T			364	T			100	99.3	96.9	BC	-
467	E			365	E			100	99.3	99.7	BC	-
468	W	S	12	366	W			40	92.1	M at 73.3	BC	VUS
469	I			367	I			65	98.6	V at 60.1	BC	-
470	A	T	12	368	A			80	98.6	46.2	BC	VUS
471	E			369	E			40	93.5	G at 46.2	BC	-
472	V	I	52	370	I			70	I at 56.1	83.7	BC	LND
473	N			371	N			100	98.6	98.5	BC	-
474	L	F	1	372	L			100	97.8	94.5	BC	PD
475	P	L	1	373	P			100	99.3	99.7	BC	LND
476	A			374	A			100	98.6	97.8	BC	-
477	A			375	A			90	89.9	85.9	BC	-
478	Q	K	28	376	Q			100	99.3	97.6	BC	LND
479	V			377	V			45	95.7	L at 78.4	BC	-
480	A			378	A			45	85.6	Q at 75.3	BC	-
481	V			379	V			55	90.7	I at 66.4	BC	-
482	G			380	G			50	97.1	A at 72.6	BC	-
483	M			381	M			100	99.3	99.7	BC	-

484	G			382	G			100	99.3	99.7	BC	-
485	I			383	I			85	92.8	75.1	BC	-
486	P			384	P			100	99.3	99.3	BC	-
487	L			385	L			90	99.3	90.1	BC	-
488	W			386	W			35	83.5	H at 57.0	BC	-
489	Q			387	Q			35	79.1	R at 68.7	BC	-
490	I			388	I			80	83.5	84.6	BC	-
491	P			389	P			50	94.2	R at 33.1	BC	-
492	E			390	E			45	97.8	73	BC	-
493	I			391	I	L	45	100	99.3	94.6	BC	ND
494	R	G	1	392	R			100	99.3	99.9	BC	PD
495	R			393	R			45	98.6	L at 37.3	BC	-
496	F			394	F			45	97.1	L at 63.0	BC	-
497	Y			395	Y			95	97.1	97	BC	-
498	G	A	1	396	G			100	98.6	95.4	BC	VUS
499	M			397	I			35	64.8	V at 36.1	BC	LND
500	E			398	E			30	60.4	S at 20.2	BC	-
501	H			399	H			30	65.5	13.8	BC	-
502	G			400	G			40	93.5	19.6	BC	-
503	G			401	G			85	87.1	D at 34.5	BC	-
504	G			402	G			40	90.7	19	BC	-
505	Y			403	Y			40	87.8	18.4	BC	-
506	D			404	D			35	83.5	17.5	BC	-
507	S			405	S			20	A at 48.9	P at 71.4	BC	-
508	W			406	W			45	96.4	48.9	BC	-
509	R			407	R			40	92	17.2	BC	-
510	K			408	K			30	74.1	15.6	BC	-
511	T			409	T			45	79.9	16.8	BC	-
512	S			410	S			45	68.4	14.4	BC	-
513	V			411	V			15	A at 47.5	A at 10.2	BC	-
514	V	L	1	412	V	A	84	20	L at 49.6	G at 40.0	BC	ND
515	A			413	A			45	95.7	29.8	BC	-
516	S			414	F			15	T at 82.0	T at 42.7	BC	LND
517	P			415	P	L	6	35	76.3	43.8	BC	ND
518	F			416	F			45	98.6	I at 76.6	BC	-
519	D			417	D			30	69.8	82.3	BC	-
520	F	L	2	418	F			40	82.7	94	BC	PD
521	D	N	52	419	D			45	95.7	43.2	BC	VUS
522	E			420	K			15	K at 57.6	K at 32.1	BC	LND
523	A			421	A			45	69.8	T at 19.0	BC	-
524	E			422	Q			15	41	Q at 44.8	BC	LND

525	S			423	S			45	97.1	R at 40.0	BC	-
526	L			424	I			5	T at 30.9	R at 33.1	BC	LND
527	R			425	R			30	48.2	P at 73.0	BC	-
528	P			426	P	S (A)	32 (1)	90	99.3	90.1	BC	ND
529	K			427	K			65	95.7	61.5	BC	-
530	G			428	G			100	99.3	95.4	BC	-
531	H			429	H			100	99.3	99.4	BC	-
532	C			430	C			60	94.2	V at 41.7	BC	-
533	V			431	V			50	96.4	I at 42.6	BC	-
534	A			432	A			95	99.3	97.8	BC	-
535	V			433	V			45	98.6	C at 37.2	BC	-
536	R			434	R			100	99.3	99.9	BC	-
537	V			435	V			35	71.9	I at 81.3	BC	-
538	T	A	1	436	T			100	99.3	99.9	BC	PD
539	S			437	S			100	98.6	84.6	BC	-
540	E			438	E			100	99.3	99.9	BC	-
541	D			439	D			65	91.4	57.7	BC	-
542	P			440	P			100	98.6	96.7	BC	-
543	D			441	D			80	99.3	62.1	BC	-
544	D			442	D			50	98.6	E at 65.1	BC	-
545	G			443	G			100	99.3	98.1	BC	-
546	F			444	F			100	99.3	99.4	BC	-
547	K			445	K			100	97.8	93.1	BC	-
548	P			446	P			100	99.3	99.7	BC	-
549	T			447	T			45	98.6	S at 69.6	BC	-
550	S			448	S			65	65.5	67.8	BC	-
551	G			449	G			100	99.3	99.3	BC	-
552	E			450	R			5	K at 74.1	T at 47.7	BC	LND
553	I			451	V			15	95	V at 53.2	BC	LND
554	Q	H	1	452	Q			60	55.4	46.5	BC	VUS
555	E			453	E			90	98.6	89.2	BC	-
556	L			454	L			90	69.8	89.4	BC	-
557	S			455	S			35	79.1	N at 76.3	BC	-
558	F			456	F			100	98.6	99.3	BC	-
559	K			457	K			45	95	R at 74.7	BC	-
560	S			458	S			100	97.1	95.5	BC	-
561	K	Q (N)	36 (2)	459	K			40	93.5	S at 57.1	BC	VUS
562	P			460	P			45	95	S at 34.9	BC	-
563	N			461	N			100	97.1	90.3	BC	-
564	M			462	V			5	V at 93.5	V at 93.3	BC	LND
565	W			463	W			100	97.1	99.3	BC	-

566	S			464	A			5	A at 85.6	G at 80.2	BC	LND
567	Y			465	Y			100	97.1	98.9	BC	-
568	F			466	F			100	97.1	99.1	BC	-
569	S			467	S			100	97.1	98.4	BC	-
570	V			468	V			100	97.1	97.8	BC	-
571	K			469	K			45	96.4	G at 39.3	BC	-
572	S			470	S			50	97.8	35.4	BC	-
573	G			471	G			45	97.8	A at 36.9	BC	-
574	G			472	G			100	98.6	92.7	BC	-
575	G			473	G			80	84.9	78.4	BC	-
576	I			474	I			65	97.8	58.3	BC	-
577	H			475	H			100	98.6	99.3	BC	-
578	E			476	E			85	97.8	62.2	BC	-
579	F			477	F			100	97.1	91.5	BC	-
580	S			478	S			50	70.5	A at 49.9	BC	-
581	D			479	D			100	98.6	99	BC	-
582	S			480	S			100	98.6	99.1	BC	-
583	Q			481	Q			100	98.6	99.1	BC	-
584	F			482	F			100	98.6	97.5	BC	-
585	G			483	G			100	98.6	99.1	BC	-
586	H			484	H			100	100	99.4	BC	-
587	V			485	V			40	96.4	I at 45.3	BC	-
588	F			486	F			100	100	99.6	BC	-
589	A	S	1	487	A			60	100	66.7	BC	VUS
590	F			488	F			45	71.9	Y at 34.2	BC	-
591	G			489	G			95	100	96.9	BC	-
592	E			490	E			85	75.5	81.3	BC	-
593	S			491	S			50	86.3	N at 58.9	BC	-
594	R			492	R			100	100	99.3	BC	-
595	S			493	A			15	A at 56.8	E at 31.6	BC	LND
596	V			494	L			5	L at 61.2	A at 37.3	BC	LND
597	A			495	A			85	99.3	63.7	BC	-
598	I			496	I			75	99.3	R at 54.1	BC	-
599	A			497	A			35	70.5	K at 43.6	BC	-
600	N			498	N			85	89.9	57.4	BC	-
601	M			499	M			95	100	88.5	BC	-
602	V			500	V			90	71.2	87	BC	-
603	L			501	L			50	99.3	V at 51.1	BC	-
604	A			502	G			75	G at 66.2	82.5	BC	LND
605	L			503	L			95	100	97.6	BC	-
606	K			504	K			100	100	96.4	BC	-

607	E			505	E			95	100	93.7	BC	-
608	I			506	I			45	92.8	L at 73.6	BC	-
609	Q			507	Q			45	96.4	S at 72.6	BC	-
610	I			508	I			100	100	96.7	BC	-
611	R			509	R			100	100	98.5	BC	-
612	G			510	G			100	100	98.1	BC	-
613	D			511	E			55	E at 97.8	70.9	BC	LND
614	I			512	I			40	95.7	F at 71.2	BC	-
615	R			513	R			85	66.2	85.2	BC	-
616	T			514	T			90	77.7	93	BC	-
617	N			515	N			45	99.3	T at 71.8	BC	-
618	V	I	9	516	V			95	100	90.9	BC	VUS
619	D			517	D			45	99.3	E at 74.2	BC	-
620	Y			518	Y			100	99.3	98.8	BC	-
621	T			519	T			45	94.2	L at 73.6	BC	-
622	I			520	I			70	52.5	77.4	BC	-
623	D			521	D			45	97.1	K at 54.9	BC	-
624	L			522	L			100	98.6	97	BC	-
625	L			523	L			100	100	96.3	BC	-
626	H			524	H			20	N at 55.4	E at 75.9	BC	-
627	A			525	A			45	98.6	T at 67.9	BC	-
628	S			526	S			30	66.2	E at 41.2	BC	-
629	D			527	D			45	74.8	33.1	BC	-
630	Y			528	Y			35	71.9	F at 83.7	BC	-
631	R	W (Q)	2 (1)	529	R			45	86.3	E at 37.8	BC	VUS
632	E			530	D			45	75.5	D at 32.5	BC	LND
633	N			531	N			100	100	96.9	BC	-
634	K			532	K			40	83.5	T at 41.8	BC	-
635	I			533	I			95	98.6	92.4	BC	-
636	H			534	H			45	99.3	T at 34.9	BC	-
637	T			535	T			100	100	99.7	BC	-
638	G			536	G			95	99.3	80.7	BC	-
639	W			537	W			100	100	99.6	BC	-
640	L			538	L			100	100	99.7	BC	-
641	D			539	D			100	100	99.6	BC	-
642	S			540	S			35	69.8	E at 22.0	BC	-
643	R			541	R			45	99.3	L at 73.0	BC	-
644	I			542	I			100	100	99.3	BC	-
645	A			543	A			80	100	50.2	BC	-
646	M			544	M			45	99.3	E at 28.0	-	-
647	R			545	R			55	97.8	K at 55.3	-	-

648	V			546	V			80	97.8	47.2	-	-
649	R			547	R			35	71.2	T at 38.7	-	-
650	A			548	A			85	97.1	79	-	-
651	E			549	E			90	98.6	84.6	-	-
652	R			550	R			75	99.3	69.9	-	-
653	P			551	P			95	97.8	92.4	-	-
654	P			552	P			45	95.7	D at 65.8	-	-
655	W			553	W			45	99.3	T at 27.0	-	-
656	Y			554	Y			45	94.2	M at 38.7	-	-
657	L	I	1	555	L			90	69.1	70.3	-	VUS
658	S			556	S			45	99.3	A at 45.6	-	-
659	V			557	V			95	100	85.3	-	-
660	V			558	V			85	96.4	59.1	-	-
661	G			559	G			45	91.4	C at 65.5	-	-
662	G			560	G			100	99.3	92.1	-	-
663	A			561	A			95	93.5	92.1	-	-
664	L			562	L			85	97.8	50.7	-	-
665	Y			563	Y			45	92.8	T at 33.6	-	-
666	K			564	K			50	82.7	55.8	-	-
667	A			565	A			90	70.5	89.4	-	-
668	S			566	S			35	68.4	H at 33.0	-	-
669	T			567	A			15	A at 41.7	A at 18.6	-	LND
670	T			568	T			25	S at 47.5	A at 36.9	-	-
671	S			569	S			40	61.9	48	-	-
672	S			570	A			15	A at 72.7	E at 36.9	-	LND
673	A			571	A			30	69.1	40.6	-	-
674	V			572	V			35	32.4	C at 28.8	-	-
675	V	G	1	573	V			50	98.6	31.2	-	VUS
676	S			574	S			60	82	40.2	-	-
677	D			575	D			45	71.2	E at 42.1	-	-
678	Y			576	Y			60	100	66.4	-	-
679	V			577	V			40	77.7	L at 23.1	-	-
680	G			578	G			30	71.2	H at 19.9	-	-
681	Y			579	Y			50	99.3	S at 40.6	-	-
682	L			580	L			100	100	90.1	-	-
683	E			581	E			65	51.1	67.9	-	-
684	K			582	K			60	100	61.6	-	-
685	G			583	G			100	100	99.4	-	-
686	Q			584	Q			95	99.3	96.7	-	-
687	I			585	I			50	98.6	V at 65.1	-	-
688	P			586	P			60	100	64	-	-

689	P			587	P			80	98.6	46.8	-	-
690	K			588	K			60	99.3	55.6	-	-
691	H	Q	1	589	H			60	95	D at 35.4	-	VUS
692	I			590	I			50	98.6	30.7	-	-
693	S			591	S			45	99.3	L at 73.6	-	-
694	L			592	L			70	97.1	40	-	-
695	V			593	V			45	97.8	T at 42.3	-	-
696	H			594	H			30	N at 45.3	V at 31.6	-	-
697	S			595	S			35	76.3	F at 41.1	-	-
698	Q			596	Q			25	48.2	P at 28.5	-	-
699	V			597	V	M	1	95	89.2	68.5	-	LND
700	S			598	S			30	69.1	E at 40.9	-	-
701	L	M	1	599	L			80	99.3	53.4	-	VUS
702	N			600	N			40	97.1	I at 71.7	-	-
703	I			601	I			45	99.3	Y at 58.6	-	-
704	E			602	E			65	85.6	70.2	-	-
705	G	E (R)	2 (1)	603	G			85	84.9	77.8	-	VUS
706	S			604	S	N	1	45	89.9	22.3	-	LND
707	K			605	K			75	99.3	58	-	-
708	Y			606	Y			100	100	96	-	-
709	T			607	T			45	92.1	K at 38.4	-	-
710	I			608	I			45	87.8	F at 43.2	-	-
711	D			609	D			25	E at 48.9	T at 39.9	-	-
712	V			610	V			65	M at 39.6	41.8	-	-
713	V			611	V			35	74.8	T at 45.1	-	-
714	R			612	R			85	89.9	73.3	-	-
715	G			613	G			30	58.3	S at 44.1	-	-
716	G			614	G			55	96.4	S at 55.2	-	-
717	S			615	S			20	P at 47.5	P at 36.7	-	-
718	G			616	G			40	83.5	D at 33.0	-	-
719	T			617	T			10	S at 86.3	S at 58.2	-	-
720	Y			618	Y			85	97.1	78	-	-
721	R			619	R			25	70.5	V at 22.3	-	-
722	L			620	L			80	97.8	75.4	-	-
723	R			621	R			40	85.6	F at 32.4	-	-
724	M			622	M			70	80.6	46.2	-	-
725	N	S	44	623	N			100	95.7	96.9	-	LND
726	N			624	K	N	3	15	E at 25.9	G at 66.3	-	ND
727	S			625	S			90	97.1	79	-	-
728	E			626	E			35	67.6	K at 28.3	-	-
729	V			627	V			45	I at 56.1	C at 31.6	-	-

730	V			628	V			15	E at 81.3	E at 53.1	-	-
731	A			629	A			45	91.4	V at 57.9	-	-
732	E			630	E			40	69.8	G at 31.8	-	-
733	I			631	I			35	65.5	V at 54.7	-	-
734	H			632	H			70	69.8	R at 44.2	-	-
735	T			633	T			30	71.9	R at 28.0	-	-
736	L			634	L			100	97.1	94.2	-	-
737	R	G	12	635	R			35	69.1	S at 48.1	-	VUS
738	D			636	D			95	97.8	94.3	-	-
739	G	E	1	637	G			95	97.8	95.2	-	PD
740	G			638	G			95	92.1	90	-	-
741	L			639	L			95	89.2	91.3	-	-
742	L			640	L	S	1	100	91.4	95.1	-	VUS
743	M	I	5	641	M			45	85.6	L at 28.3	-	VUS
744	Q			642	Q			45	89.9	S at 32.8	-	-
745	L			643	L			60	95.7	58	-	-
746	D			644	D			65	93.5	58.5	-	-
747	G			645	G			100	95	95.5	-	-
748	K			646	K			30	83.5	25	-	-
749	S			647	S			95	95.7	91.6	-	-
750	H			648	H			60	95.7	66.9	-	-
751	V			649	V			45	88.5	T at 39.3	-	-
752	I			650	I			50	87.8	V at 34.0	-	-
753	Y			651	Y			100	95.7	95.5	-	-
754	A			652	A			45	93.5	W at 37.6	-	-
755	K			653	E			55	E at 94.2	60	-	LND
756	E			654	E			80	83.5	86.2	-	-
757	E			655	E			95	99.3	94.3	-	-
758	A			656	A			40	94.2	V at 52.8	-	-
759	T			657	A			10	A at 67.6	A at 29.1	-	LND
760	G			658	G			55	98.6	A at 35.8	-	-
761	T			659	T			65	98.6	55.3	-	-
762	R	C	6	660	R			100	97.1	96.6	-	ND
763	L			661	L			55	98.6	60.6	-	-
764	L			662	L			40	95	S at 37.0	-	-
765	I			663	I			85	98.6	57.7	-	-
766	D			664	D	G	1	55	82.7	62.1	-	LND
767	G	R	1	665	G			45	98.6	43.2	-	VUS
768	R	S	1	666	R			30	67.6	62.2	-	VUS
769	T			667	T			95	99.3	94.6	-	-
770	C	F	1	668	C			90	99.3	83.7	-	VUS

771	L			669	L			55	95	59.1	-	-
772	L			670	L			65	98.6	56.7	-	-
773	Q			671	Q			45	98.6	E at 69.1	-	-
774	N	S	1	672	N			40	82.7	K at 35.7	-	VUS
775	D	V	1	673	D			30	81.3	E at 75.0	-	VUS
776	H	D (N)	3 (1)	674	H			45	98.6	N at 61.6	-	VUS
777	D	N	4	675	D			95	100	97.2	-	PD
778	P			676	P			100	100	99	BCCP	-
779	S			677	S			65	99.3	T at 55.3	BCCP	-
780	K			678	K			35	73.4	Q at 40.9	BCCP	-
781	L			679	L			85	100	88.9	BCCP	-
782	M			680	M			20	L at 41.0	R at 72.0	BCCP	-
783	A	V	1	681	A			40	97.1	S at 43.2	BCCP	VUS
784	E			682	E			35	83.5	P at 70.2	BCCP	-
785	T			683	T			50	99.3	S at 68.7	BCCP	-
786	P			684	P			65	100	65.2	BCCP	-
787	C			685	C			45	97.1	G at 79.0	BCCP	-
788	K	M	1	686	K			100	99.3	97.2	BCCP	VUS
789	L			687	L			100	100	96.3	BCCP	-
790	L			688	M			45	89.2	V at 45.7	BCCP	LND
791	R			689	R			50	100	36.9	BCCP	-
792	Y			690	Y			65	F at 56.8	55.5	BCCP	-
793	L			691	L			60	99.3	54.9	BCCP	-
794	V			692	V	I	37	95	84.9	85.6	BCCP	ND
795	S			693	S	F	1	20	A at 54.7	E at 60.4	BCCP	LND
796	D			694	D	H	1	80	99.3	56.5	BCCP	LND
797	N			695	N			15	G at 71.9	G at 91.2	BCCP	-
798	S			696	S			40	68.4	G at 21.7	BCCP	-
799	S			697	N			15	H at 84.9	H at 90.0	BCCP	LND
800	I	M	1	698	I			35	V at 70.5	V at 72.7	BCCP	VUS
801	D			699	D			35	77.7	20.5	BCCP	-
802	T			700	A			5	92.1	A at 71.4	BCCP	LND
803	D			701	D			45	98.6	G at 76.9	BCCP	-
804	T			702	T	M	2	30	61.2	Q at 51.7	BCCP	ND
805	P	R	2	703	P			55	98.6	43	BCCP	VUS
806	Y			704	Y			95	100	69.1	BCCP	-
807	A			705	A			100	97.8	98.4	BCCP	-
808	E			706	E			100	100	97.8	BCCP	-
809	V			707	V	A	1	55	98.6	52.3	BCCP	LND
810	E			708	E			100	100	99.9	BCCP	-
811	V			709	V			95	100	94.5	BCCP	-

812	M			710	M			100	100	100	BCCP	-
813	K			711	K			100	100	100	BCCP*	-
814	M			712	M			100	100	97.9	BCCP	-
815	C			713	C			45	99.3	Y at 29.8	BCCP	-
816	M			714	M			100	99.3	93	BCCP	-
817	P			715	P			65	100	67.5	BCCP	-
818	L			716	L			100	99.3	95.4	BCCP	-
819	I			717	L			10	L at 97.8	L at 25.3	BCCP	LND
820	S			718	S			45	79.1	A at 55.8	BCCP	-
821	P			719	P			45	99.3	Q at 37.9	BCCP	-
822	A			720	A	S	1	45	97.1	E at 70.3	BCCP	LND
823	S			721	S			65	84.9	45.4	BCCP	-
824	G			722	G			100	100	100	BCCP	-
825	V			723	V			55	74.1	34.3	BCCP	-
826	I			724	I			65	95.7	V at 47.5	BCCP	-
827	H			725	H			55	69.8	Q at 37.5	BCCP	-
828	F			726	F			45	69.1	33	BCCP	-
829	K			727	K			20	43.2	I at 31.8	BCCP	-
830	L			728	M			10	M at 76.3	K at 67.8	BCCP	LND
831	S			729	S			25	75.5	Q at 40.0	BCCP	-
832	E			730	E			45	97.1	P at 65.5	BCCP	-
833	G	R	3	731	G			100	98.6	99.3	BCCP	PD
834	Q			732	Q			45	97.8	A at 49.9	BCCP	-
835	A			733	A			45	84.2	T at 35.8	BCCP	-
836	M			734	M			40	87.1	L at 64.9	BCCP	-
837	Q			735	Q			35	92.8	E at 44.5	BCCP	-
838	A	V	1	736	A			75	93.5	68.8	BCCP	VUS
839	G			737	G			90	84.9	96.7	BCCP	-
840	E			738	E			25	D at 50.4	D at 54.9	BCCP	-
841	L			739	L			50	100	I at 42.6	BCCP	-
842	I			740	I			55	99.3	45.6	BCCP	-
843	A			741	A			80	99.3	64.5	BCCP	-
844	K			742	N			35	R at 71.2	I at 37.0	BCCP	LND
845	L			743	L			80	100	85	BCCP	-
846	D			744	D			50	96.4	A at 24.0	CEN	-
847	L	P	1	745	L			100	100	96	CEN	PD
848	D			746	D			95	96.4	94.8	CEN	-
849	D			747	D			85	100	88.9	CEN	-
850	P			748	P			100	97.8	97	CEN	-
851	S			749	S			100	99.3	90.7	CEN	-
852	A			750	A			30	81.3	K at 32.4	CEN	-

853	V			751	V			90	100	93.6	CEN	-
854	R			752	R			35	69.1	K at 46.6	CEN	-
855	K			753	K			30	53.2	H at 25.0	CEN	-
856	A			754	A			90	97.1	86.4	CEN	-
857	K	E	1	755	E			10	E at 95.0	E at 39.1	CEN	LND
858	P	A	9	756	P	L	3	80	97.8	64.8	CEN	ND
859	F	L	1	757	F			75	100	74.2	CEN	VUS
860	R	H	6	758	H			10	H at 49.6	T at 31.9	CEN	LND
861	G			759	G			95	95	84.4	CEN	-
862	S			760	S			35	55.4	Q at 25.3	CEN	-
863	F			761	F			50	99.3	L at 64.6	CEN	-
864	P			762	P			95	100	91.3	CEN	-
865	R			763	R			25	V at 23.7	E at 16.0	CEN	-
866	L	F	1	764	L			40	71.9	32.7	CEN	VUS
867	G	E	1	765	G			55	74.8	56.8	CEN	VUS
868	L			766	L			30	P at 55.4	P at 29.5	CEN	-
869	P			767	P			65	98.6	62.4	CEN	-
870	T	R (P)	2 (2)	768	T			35	69.8	17.1	CEN	VUS
871	A			769	A			40	93.5	I at 24.4	CEN	-
872	I			770	I			35	54	V at 24.7	CEN	-
873	S			771	S			45	94.2	G at 63.7	CEN	-
874	G			772	G			35	77	E at 28.2	CEN	-
875	K	E	1	773	R	K	259	75	71.9	84	CEN	ND
876	V	I	1	774	V			45	98.6	P at 41.2	CEN	VUS
877	H			775	H			75	100	50.2	CEN	-
878	Q			776	Q			60	69.8	56.1	CEN	-
879	R			777	R			40	69.1	44.8	CEN	-
880	C			778	C			30	74.1	F at 58.2	CEN	-
881	A			779	A			45	97.8	26.4	CEN	-
882	A			780	A			45	88.5	22.5	CEN	-
883	T			781	T			25	S at 84.9	L at 33.6	CEN	-
884	L			782	L			70	76.3	46.8	CEN	-
885	N			783	N			45	91.4	32.5	CEN	-
886	A			784	A			40	81.3	I at 26.4	CEN	-
887	A			785	A			40	92.1	L at 71.4	CEN	-
888	R			786	R	C	1	35	73.4	E at 19.0	CEN	LND
889	M	L	21	787	M			45	97.8	N at 53.7	CEN	VUS
890	I			788	I			60	71.9	61.8	CEN	-
891	L			789	L			70	100	72.6	CEN	-
892	A			790	A			50	97.8	31.5	CEN	-
893	G			791	G			100	99.3	97.8	CEN	-

894	Y			792	Y			95	99.3	79.3	CEN	-
895	D			793	E			25	E at 71.9	45.7	CEN	LND
896	H	D (N)	12 (1)	794	H			45	97.8	L at 28.0	CEN	VUS
897	K	R	2	795	K			40	N at 68.4	19.6	CEN	VUS
898	V			796	V			30	I at 69.8	M at 37.8	CEN	-
899	D			797	D			25	53.2	K at 23.8	CEN	-
900	E			798	E			45	64.8	28.2	CEN	-
901	V			799	V			45	97.1	T at 30.1	CEN	-
902	L			800	V			20	V at 93.5	V at 44.7	CEN	LND
903	Q	H	78	801	Q			60	82	29.5	CEN	VUS
904	D			802	D			45	55.4	E at 25.2	CEN	-
905	L			803	L			90	97.8	84.4	CEN	-
906	L			804	L			30	67.6	M at 28.2	CEN	-
907	N	H	1	805	N			30	48.2	E at 26.5	CEN	VUS
908	C			806	C			45	97.8	V at 36.6	CEN	-
909	L			807	L			100	98.6	93.4	CEN	-
910	D			808	D			45	97.1	R at 60.9	CEN	-
911	S	T	8	809	S			30	67.8	D at 51.1	CEN	VUS
912	P			810	P			90	99.3	85.5	CEN	-
913	E			811	E			50	90.7	46.2	CEN	-
914	L			812	L			100	99.3	95.8	CEN	-
915	P			813	P			100	98.6	95.4	CEN	-
916	F			814	F			45	95	Y at 37.9	CEN	-
917	L			815	L			80	97.8	50.5	CEN	-
918	Q	L	1	816	Q			45	97.8	E at 68.5	CEN	VUS
919	W			817	W			60	98.6	53.1	CEN	-
920	Q			818	Q			70	71.2	45.7	CEN	-
921	E			819	E			65	98.6	39.3	CEN	-
922	C			820	C			25	51.8	I at 25.5	CEN	-
923	F			821	F			30	M at 53.2	M at 35.2	CEN	-
924	A			822	A			30	S at 56.1	S at 54.3	CEN	-
925	V			823	V			45	97.8	A at 34.3	CEN	-
926	L			824	L			60	99.3	63.1	CEN	-
927	A			825	A			60	97.1	H at 31.2	CEN	-
928	T			826	T			45	87.1	G at 37.3	CEN	-
929	R			827	R			100	100	93.1	CEN	-
930	L			828	L			55	95.7	I at 36.1	CEN	-
931	P			829	P			95	100	94.5	CEN	-
932	K			830	K			35	68.4	Q at 22.9	CEN	-
933	D			831	N			30	61.2	K at 33.4	CEN	LND
934	L			832	L			65	100	63.4	CEN	-

935	R	K	2	833	R			35	50.4	E at 37.6	CEN	VUS
936	N			834	N			25	52.5	K at 28.2	CEN	-
937	M			835	M	I	6	20	E at 79.1	Q at 21.7	CEN	ND
938	L			836	L			55	97.8	41.4	CEN	-
939	E			837	E			40	76.3	R at 24.9	CEN	-
940	L			838	S			5	S at 33.8	K at 22.0	CEN	LND
941	K	I	1	839	K			70	86.3	L at 19.9	CEN	VUS
942	Y			840	Y			45	84.2	M at 30.4	CEN	-
943	K			841	R			25	57.6	A at 24.6	CEN	LND
944	E			842	E			55	59.2	R at 26.5	CEN	-
945	F			843	F			20	Y at 58.3	Y at 42.6	CEN	-
946	E			844	E			35	70.5	A at 24.6	CEN	-
947	I			845	S			10	L at 42.5	S at 40.6	CEN	LND
948	I			846	I			25	46	N at 36.7	CEN	-
949	S	F	109	847	S			25	49.6	I at 28.5	CEN	VUS
950	K			848	R			10	S at 30.9	T at 28.9	CEN	LND
951	T			849	N			40	H at 22.3	S at 38.1	CEN	LND
952	S			850	S			60	35.3	V at 27.7	CEN	-
953	L			851	L	M	1	20	K at 42.5	K at 8.9	CEN	LND
954	T			852	T			25	N at 39.6	L at 29.4	CEN	-
955	P			853	T			5	K at 36.7	C at 21.3	CEN	LND
956	D			854	D			45	89.2	E at 29.8	CEN	-
957	F			855	F			100	100	94.8	CEN	-
958	P			856	P			100	100	91.9	CEN	-
959	A			857	A			55	75.5	53.4	CEN	-
960	K			858	K			50	82.7	38.4	CEN	-
961	L	V	1	859	L			40	68.4	Q at 46.0	CEN	VUS
962	L			860	L			65	97.8	55.8	CEN	-
963	K			861	K			25	R at 62.6	A at 30.3	CEN	-
964	G	R	9	862	G			30	62.6	K at 37.8	CEN	VUS
965	I			863	I			65	51.1	33.6	CEN	-
966	L			864	L			50	51.1	43.8	CEN	-
967	E			865	E			45	94.2	D at 38.7	CEN	-
968	A			866	A			35	67.6	S at 22.2	CEN	-
969	H			867	H			60	52.5	40	CEN	-
970	L			868	L			55	93.5	35.4	CEN	-
971	S			869	S			20	A at 43.9	A at 40.6	CEN	-
972	S			870	S			25	48.9	T at 29.5	CEN	-
973	C			871	C			35	66.2	L at 31.6	CEN	-
974	D			872	D			20	S at 48.9	V at 22.2	CEN	-
975	E	K	192	873	E			45	74.1	R at 19.0	CEN	VUS

976	K			874	K			35	93.5	40.9	CEN	-
977	E			875	E			45	79.1	30.9	CEN	-
978	R			876	R			65	K at 43.2	38.8	CEN	-
979	G			877	G			25	48.2	E at 22.6	CEN	-
980	S	A	9	878	A			5	T at 48.9	V at 26.1	CEN	LND
981	L			879	L			40	Q at 33.1	F at 36.3	CEN	-
982	E			880	E			45	95	F at 28.9	CEN	-
983	R	S	13	881	R			45	99.3	M at 21.4	CEN	VUS
984	L			882	L			45	97.8	T at 26.2	CEN	-
985	I			883	I			25	V at 72.7	T at 28.9	CEN	-
986	E			884	E			50	90.7	Q at 31.8	CEN	-
987	P			885	P			65	100	62.1	CEN	-
988	L			886	L			65	100	64	CEN	-
989	M			887	M			40	78.4	V at 34.2	CEN	-
990	S			888	S	N	7	40	89.2	Q at 37.2	CEN	ND
991	L			889	L			80	95.7	52	CEN	-
992	V			890	A			55	59	41.1	CEN	LND
993	K			891	K			45	96.4	Q at 28.9	CEN	-
994	S			892	S			45	98.6	R at 43.6	CEN	-
995	Y	F	1	893	Y			100	99.3	89.1	CEN	VUS
996	E			894	E			45	100	R at 30.6	CEN	-
997	G			895	G			45	97.8	25.2	CEN	-
998	G	S	1	896	G			100	100	95.4	CEN	VUS
999	R			897	R			45	98.6	L at 32.1	CEN	-
1000	E			898	E			50	97.8	R at 31.2	CEN	-
1001	S			899	S			45	89.9	G at 33.4	CEN	-
1002	H			900	H			80	100	66.4	CEN	-
1003	A	T	1	901	A			45	97.8	E at 40.8	CEN	VUS
1004	R	H	1	902	R			25	56.8	K at 27.7	CEN	VUS
1005	L			903	V			5	V at 36.7	A at 23.4	CEN	LND
1006	I			904	I			40	71.2	V at 57.1	CEN	-
1007	V			905	V			80	98.6	51.7	CEN	-
1008	H			906	H			20	K at 43.9	K at 13.2	CEN	-
1009	S			907	S			45	94.2	30.6	CEN	-
1010	L			908	L			95	100	85.3	CEN	-
1011	F			909	F			45	98.6	L at 65.4	CEN	-
1012	E			910	E			55	90.7	42.6	CEN	-
1013	E			911	E			50	89.9	Q at 33.9	CEN	-
1014	Y			912	Y			100	100	96	CEN	-
1015	L			913	L			75	99.3	45.3	CEN	-
1016	S			914	S			40	80.6	22.2	CEN	-

1017	V			915	V			95	95	87.1	CEN	-
1018	E			916	E			100	100	99.3	CEN	-
1019	E			917	E			45	98.6	22.6	CEN	-
1020	L			918	L			60	95.7	55.2	CEN	-
1021	F			919	F			100	100	97.2	CEN	-
1022	N			920	N			30	S at 79.9	S at 32.5	CEN	-
1023	D			921	D			45	94.2	G at 25.6	CEN	-
1024	N			922	N			30	55.4	H at 28.6	CEN	-
1025	M			923	M			20	I at 82.0	Y at 28.3	CEN	-
1026	L			924	L			20	Q at 86.3	D at 28.2	CEN	-
1027	A			925	A			20	55.4	K at 27.3	CEN	-
1028	D			926	D			50	97.1	38.2	CEN	-
1029	V			927	V			55	99.3	54.1	CEN	-
1030	I	T	2	928	I			60	99.3	54.1	CEN	VUS
1031	E			929	E			45	99.3	L at 41.1	CEN	-
1032	R			930	R			45	84.2	K at 25.9	CEN	-
1033	M			931	M			20	L at 88.5	L at 88.5	CEN	-
1034	R			932	R			100	99.3	93.4	CEN	-
1035	Q			933	Q			20	L at 66.2	E at 37.3	CEN	-
1036	Q			934	L			50	91.4	39.6	CEN	LND
1037	Y			935	Y			25	55.4	N at 42.1	CEN	-
1038	K			936	K			65	51.2	69.6	CEN	-
1039	K			937	K			50	99.3	D at 35.5	CEN	-
1040	D			938	D			95	100	80.7	CEN	-
1041	R	Q	13	939	L			10	L at 95.7	L at 44.4	CEN	LND
1042	L			940	L			20	49.6	D at 14.4	CEN	-
1043	K			941	K			65	97.1	51.6	CEN	-
1044	I			942	I			30	V at 69.1	V at 85.8	CEN	-
1045	V			943	V			50	96.4	49.8	CEN	-
1046	D	Y	1	944	D			45	82.7	26.1	CEN	VUS
1047	I			945	I			50	97.1	35.5	CEN	-
1048	V			946	V			55	99.3	54.3	CEN	-
1049	L			947	L			55	84.9	59.2	CEN	-
1050	S			948	S			100	100	93.9	CEN	-
1051	H			949	H			100	98.6	97.3	CEN	-
1052	Q			950	Q			45	99.3	S at 33.7	CEN	-
1053	G			951	G			45	96.4	Q at 29.2	CEN	-
1054	I			952	I			20	V at 74.8	V at 70.9	CEN	-
1055	I			953	K			5	K at 49.6	S at 17.4	CEN	LND
1056	H			954	N			5	N at 43.2	K at 28.2	CEN	LND
1057	K			955	K			100	100	93.9	CEN	-

1058	N			956	N			95	72.7	79.5	CEN	-
1059	K			957	K			45	95.7	27.9	CEN	-
1060	L			958	L			100	100	93.6	CEN	-
1061	V			959	V	I	1	65	I at 77.0	45.3	CEN	LND
1062	L			960	L			50	87.1	53.1	CEN	-
1063	R	Q	5	961	R			30	56.1	A at 41.8	CEN	VUS
1064	L			962	L			80	99.3	59.7	CEN	-
1065	M			963	M			45	91.4	L at 48.7	CEN	-
1066	E			964	E			40	85.6	D at 51.3	CEN	-
1067	Q			965	Q			35	K at 36.0	25.6	CEN	-
1068	L			966	L			75	95	51.9	CEN	-
1069	V			967	V			45	99.3	R at 22.5	CEN	-
1070	Y			968	Y			40	97.1	P at 33.4	CEN	-
1071	P			969	P			40	95.7	25.5	CEN	-
1072	N			970	N			45	93.5	D at 20.4	CEN	-
1073	P			971	P			45	99.3	47.4	CEN	-
1074	A			972	A			35	76.3	17.2	CEN	-
1075	A			973	A			35	92.8	L at 27.0	CEN	-
1076	Y			974	Y			45	99.3	20.8	CEN	-
1077	R			975	R			50	91.4	T at 21.0	CEN	-
1078	E			976	D			5	D at 92.8	D at 43.2	CEN	LND
1079	K	N	1	977	K			20	Q at 56.8	E at 29.5	CEN	VUS
1080	L			978	L			95	100	54.3	CEN	-
1081	I			979	I			35	79.1	A at 22.2	CEN	-
1082	R			980	R			45	100	E at 28.5	CEN	-
1083	F			981	F			45	97.8	L at 68.4	CEN	-
1084	S			982	S			40	93.5	T at 42.9	CEN	-
1085	A	E	1	983	T			15	31.7	E at 29.7	CEN	LND
1086	L			984	L			100	100	95.7	CEN	-
1087	N			985	N			50	95.7	26.7	CEN	-
1088	H			986	H			45	96.4	E at 29.1	CEN	-
1089	T			987	T			45	66.2	S at 42.1	CEN	-
1090	N			988	N			20	38.9	R at 34.8	CEN	-
1091	Y	S	1	989	Y			45	99.3	T at 30.0	CEN	VUS
1092	S			990	S			50	70.5	38.5	CEN	-
1093	Q			991	E			5	E at 55.4	K at 63.9	CEN	LND
1094	L			992	L			40	95.7	V at 72.6	CEN	-
1095	A	P	1	993	A			85	95.7	74.4	CEN	VUS
1096	L			994	L			95	98.6	93.9	CEN	-
1097	K			995	K			60	97.8	57.7	CEN	-
1098	A			996	A			95	98.6	94.5	CEN	-

1099	S	G	1	997	S			45	97.1	R at 74.2	CEN	VUS
1100	Q			998	Q			70	70.5	E at 46.3	CEN	-
1101	L			999	L			50	98.6	V at 44.5	CEN	-
1102	L			1000	L			100	96.4	89.8	CEN	-
1103	E	K	210	1001	E			40	97.8	I at 68.5	CEN	VUS
1104	Q			1002	Q			45	82	50.2	CEN	-
1105	T			1003	T			40	94.2	C at 29.4	CEN	-
1106	K			1004	K			45	99.3	H at 28.9	CEN	-
1107	R	Q (G)	5 (1)	1005	L			5	L at 99.3	L at 75.3	CEN	LND
1108	S			1006	S			45	98.6	P at 77.7	CEN	-
1109	E	D	1	1007	E			45	97.1	75.9	CEN	VUS
1110	L			1008	L			55	97.8	49.2	CEN	-
1111	R			1009	R			45	97.8	97.9	CEN	-
1112	S			1010	S			25	52.5	H at 28.0	CEN	-
1113	N			1011	N			15	S at 61.2	34.2	CEN	-
1114	I			1012	I			35	94.2	Q at 70.2	CEN	-
1115	A			1013	A			45	95.7	M at 36.4	CEN	-
1116	R			1014	R			40	96.4	E at 67.0	CEN	-
1117	S	R	1	1015	S			40	84.9	47.8	CEN	VUS
1118	L			1016	L			45	100	50.5	CEN	-
1119	S			1017	S			80	99.3	83.5	CEN	-
1120	E			1018	E			40	83.5	S at 37.2	CEN	-
1121	L	I	1	1019	L			65	99.3	V at 45.3	CEN	VUS
1122	E			1020	E			40	80.6	D at 28.3	CEN	-
1123	M			1021	M			45	98.6	45.7	CEN	-
1124	F			1022	F			40	87.8	Y at 26.7	CEN	-
1125	T			1023	T			30	84.9	S at 28.2	CEN	-
1126	E			1024	E			35	89.2	H at 27.6	CEN	-
1127	A			1025	D			5	E at 54.0	Y at 38.7	CEN	LND
1128	G			1026	G			35	70.5	E at 30.0	CEN	-
1129	E			1027	E			35	64	T at 19.8	CEN	-
1130	N			1028	N			20	35.3	7.8	CEN	-
1131	M			1029	M			20	37.4	8	CEN	-
1132	D			1030	D			30	48.9	G at 24.4	CEN	-
1133	T			1031	T			40	70.5	W at 21.4	CEN	-
1134	P	L	1	1032	P			55	69.8	D at 18.3	CEN	VUS
1135	K			1033	K			30	49.6	13	CEN	-
1136	R			1034	R			40	79.9	H at 29.5	CEN	-
1137	K			1035	K			35	76.3	R at 30.4	CEN	-
1138	S			1036	S			30	51.1	E at 27.6	CEN	-
1139	A			1037	A			35	84.9	P at 39.9	CEN	-

1140	I			1038	I			40	94.2	C at 22.3	CEN	-
1141	S	R	1	1039	N			10	N at 69.1	L at 22.9	CEN	LND
1142	E			1040	E			30	74.1	59.1	CEN	-
1143	T			1041	R			5	R at 68.4	N at 28.2	CEN	LND
1144	M			1042	I			40	96.4	L at 57.3	CEN	LND
1145	E	K	1	1043	E			40	87.1	K at 43.2	CEN	VUS
1146	N			1044	D			10	D at 84.9	E at 31.9	CEN	LND
1147	L			1045	L			95	98.6	70	CEN	-
1148	V	L	2	1046	V			50	98.6	46.5	CEN	VUS
1149	S			1047	S	C	1	30	58.3	D at 42.3	CEN	LND
1150	S			1048	A			60	A at 92.1	71.8	CEN	LND
1151	S			1049	S			15	P at 90.7	E at 29.7	CEN	-
1152	L			1050	L			45	85.6	Y at 34.5	CEN	-
1153	A			1051	A			35	70.5	T at 36.0	CEN	-
1154	V			1052	V			55	100	60.3	CEN	-
1155	E	D	3	1053	E			45	99.3	F at 66.3	CEN	VUS
1156	D			1054	D			95	100	94.9	CEN	-
1157	A			1055	A			45	99.3	V at 63.4	CEN	-
1158	L			1056	L			100	100	96.7	CEN	-
1159	V	M	9	1057	V			40	82	P at 41.8	CEN	VUS
1160	G			1058	G			25	50.4	N at 14.8	CEN	-
1161	L			1059	L			45	98.6	F at 71.4	CEN	-
1162	F			1060	F			95	100	87.9	CEN	-
1163	D			1061	D			45	100	Y at 29.7	CEN	-
1164	H			1062	H			90	72.7	81.3	CEN	-
1165	S	C	2	1063	S			65	86.3	42.1	CEN	VUS
1166	D			1064	D	E	1	65	98.6	62.7	CEN	LND
1167	H			1065	H			25	53.2	P at 31.8	CEN	-
1168	T			1066	T			45	97.8	W at 28.3	CEN	-
1169	L	F	2	1067	L			45	86.3	V at 71.1	CEN	VUS
1170	Q			1068	Q			45	97.8	21.9	CEN	-
1171	R	I	1	1069	R			35	71.9	L at 27.7	CEN	VUS
1172	R	W	5	1070	R			40	85.6	A at 70.5	CEN	VUS
1173	V	L	6	1071	V			45	98.6	A at 66.4	CEN	VUS
1174	V			1072	V			35	67.6	L at 62.2	CEN	-
1175	E			1073	E			95	97.1	90.9	CEN	-
1176	T	I	1	1074	T			45	92.8	V at 68.2	CEN	VUS
1177	Y			1075	Y			100	100	96.9	CEN	-
1178	I			1076	I			50	74.1	V at 61.2	CEN	-
1179	H			1077	R			5	R at 69.8	R at 87.6	CEN	LND
1180	R			1078	R			95	100	95.2	CEN	-

1181	L	V	1	1079	L			45	99.3	A at 57.7	CEN	VUS
1182	Y			1080	Y			100	100	98.4	CEN	-
1183	Q			1081	Q			45	99.3	R at 42.3	CEN	-
1184	P			1082	P			40	88.5	A at 61.0	CEN	-
1185	Y			1083	Y			90	71.2	88	CEN	-
1186	V	G	2	1084	V			25	L at 87.8	E at 28.6	CEN	VUS
1187	V	F (I) (G)	1 (1) (1)	1085	V			50	97.1	L at 55.3	CEN	VUS
1188	K			1086	K			45	82.7	37.2	CEN	-
1189	E			1087	D			20	G at 54.7	S at 30.4	CEN	LND
1190	S			1088	S			45	99.3	I at 31.0	CEN	-
1191	V			1089	V			20	61.2	Q at 32.5	CEN	-
1192	R			1090	R			35	71.2	Y at 32.2	CEN	-
1193	M			1091	M			40	78.4	17.7	CEN	-
1194	Q			1092	Q			60	71.2	34.9	CEN	-
1195	W			1093	W			35	75.5	L at 27.3	CEN	-
1196	H			1094	H			35	69.1	15.9	CEN	-
1197	Q			1095	R			10	R at 66.9	D at 37.5	CEN	LND
1198	S			1096	S			35	89.2	22.3	CEN	-
1199	G			1097	G			45	96.4	P at 26.1	CEN	-
1200	V	L	9	1098	L			10	L at 68.4	C at 23.2	CEN	LND
1201	I			1099	L	I	267	50	88.5	34.5	CEN	ND
1202	A			1100	A			40	94.2	V at 45.0	CEN	-
1203	S			1101	S			20	44.6	33.1	CEN	-
1204	W			1102	W			65	100	64.3	CEN	-
1205	E			1103	D	E	392	45	90.7	Q at 34.5	CEN	ND
1206	F	L	1	1104	F			100	100	96.3	CEN	LD
1207	L			1105	L			25	39.6	I at 27.7	CEN	-
1208	E			1106	E			30	86.3	D at 40.3	CEN	-
-	-	-	-	1107	E			-	-	-	-	LND
1209	H			1108	H			45	72.7	L at 22.3	CEN	-
1210	F			1109	M			5	I at 37.4	S at 30.1	CEN	LND
1211	E			1110	E			35	59	Y at 36.0	CEN	-
1212	R			1111	R			25	47.5	L at 20.4	CEN	-
1213	K			1112	K			25	46.8	V at 14.1	CEN	-
1214	N			1113	N			35	69.8	30.3	CEN	-
1215	T			1114	I			5	I at 7.2	M at 15.1	CEN	LND
1216	G			1115	G			20	14.4	< 5	CEN	-
1217	P			1116	L			5	G at 22.3	G at 17.1	CEN	LND
1218	D			1117	D			20	G at 30.2	S at 11.8	CEN	-
1219	D	G (Y)	20 (2)	1118	D	N	3	20	E at 25.9	S at 13.6	CEN	ND
1220	H			1119	H	P	37	5	D at 42.5	D at 8.6	CEN	ND

1221	E			1120	D			15	Q at 23.0	T at 7.2	CEN	LND
1222	I	M	2	1121	T			10	T at 23.0	P at 33.6	CEN	LND
1223	S			1122	S			30	30.2	5.3	CEN	-
1224	E			1123	E			20	D at 19.4	< 5	CEN	-
1225	K			1124	K			30	52.5	L at 17.8	CEN	-
1226	G			1125	G			15	P at 26.6	R at 53.5	CEN	-
1227	I			1126	L			5	L at 27.3	K at 21.3	CEN	LND
1228	V			1127	V			35	50.4	10.5	CEN	-
1229	A	V	1	1128	E	G	1	5	E at 64.8	E at 15.4	CEN	LND
1230	K			1129	K			30	52.5	11.2	CEN	-
1231	S			1130	R	H	263	20	H at 34.5	V at 11.1	CEN	ND
1232	S			1131	S			45	25.2	22.2	CEN	-
1233	K			1132	K			15	E at 60.4	E at 21.6	CEN	-
1234	R			1133	R			20	K at 55.4	Q at 22.8	CEN	-
1235	K			1134	K			25	50.4	R at 39.0	CEN	-
1236	R			1135	W			5	W at 76.3	M at 20.4	CEN	LND
1237	G			1136	G			65	98.6	97.6	CEN	-
1238	T	I	10	1137	A			10	A at 61.2	A at 36.4	CEN	LND
1239	M	I	1	1138	M			50	99.3	52.9	CEN	VUS
1240	V			1139	V			50	94.2	61.8	CEN	-
1241	I			1140	I			25	64	A at 26.1	CEN	-
1242	I			1141	I			35	69.1	F at 34.9	CEN	-
1243	K			1142	K			50	95.7	30.3	CEN	-
1244	S			1143	S			45	99.3	31	CEN	-
1245	L			1144	L			60	95	55	CEN	-
1246	Q			1145	Q			30	59	E at 37.0	CEN	-
1247	F	C	1	1146	F			55	51.1	D at 42.4	CEN	VUS
1248	L			1147	L			40	74.1	F at 30.7	CEN	-
1249	P	R	3	1148	P			25	53.2	E at 27.6	CEN	VUS
1250	S			1149	S			15	T at 41.0	E at 24.7	CEN	-
1251	I			1150	I			30	A at 48.9	20.1	CEN	-
1252	I			1151	I			45	84.9	L at 35.4	CEN	-
1253	N			1152	S	T	7	20	S at 33.1	D at 27.6	CEN	ND
1254	A			1153	A			45	87.1	E at 26.2	CEN	-
1255	S			1154	A			5	A at 94.2	A at 48.0	CEN	LND
1256	L			1155	L			50	98.6	70.6	CEN	-
1257	R			1156	R			15	K at 54.0	K at 15.3	CEN	-
1258	E	K	2	1157	E			30	73.4	21.1	CEN	VUS
1259	T			1158	T	A	1	30	59.7	S at 24.0	CEN	LND
1260	N			1159	K	M	1	10	S at 38.1	P at 19.6	CEN	LND
1261	H			1160	H			15	61.2	13	CEN	-

1262	S			1161	N			20	Y at 34.5	T at 12.0	CEN	LND
1263	H			1162	D			5	L at 24.5	S at 16.9	CEN	LND
1264	C			1163	Y	S (C)	262 (1)	5	H at 30.2	D at 11.4	CEN	ND
1265	E			1164	E			15	A at 15.8	P at 9.0	CEN	-
1266	Y	N	1	1165	T			5	V at 21.6	R at 19.6	CEN	LND
1267	A			1166	A			35	S at 29.5	< 5	CEN	-
1268	R			1167	G			10	N at 35.3	< 5	CEN	LND
1269	A			1168	A			15	G at 48.9	P at 9.6	CEN	-
1270	P			1169	P			30	S at 32.4	S at 10.6	CEN	-
1271	L			1170	L	F	3	35	18.7	S at 10.0	CEN	ND
1272	S			1171	S			25	E at 25.9	L at 33.0	CEN	-
1273	G			1172	G			40	85.6	S at 22.0	CEN	-
1274	N			1173	N			50	97.8	40.6	CEN	-
1275	M			1174	M			35	75.5	V at 40.0	CEN	-
1276	M			1175	M			35	61.9	I at 33.3	CEN	-
1277	H			1176	H			45	99.3	50.1	CEN	-
1278	I			1177	I			45	89.9	58.8	CEN	-
1279	A			1178	A			40	89.9	49	CEN	-
1280	V			1179	I			20	L at 87.7	L at 24.0	CEN	LND
1281	V			1180	V			30	67.6	44.2	CEN	-
1282	G			1181	G			35	82	A at 27.3	CEN	-
1283	I			1182	I			30	64	43.3	CEN	-
1284	N	H	1	1183	N			35	67.6	E at 29.5	CEN	VUS
1285	N			1184	N			50	86.3	20.2	CEN	-
1286	Q			1185	Q			35	75.5	17.1	CEN	-
1287	M			1186	M			35	80.6	D at 32.8	CEN	-
1288	S			1187	S			30	71.2	D at 17.5	CEN	-
1289	L			1188	L			45	47.5	16.3	CEN	-
1290	L			1189	L			30	66.9	E at 16.6	CEN	-
1291	Q			1190	Q			30	71.9	16.5	CEN	-
1292	D			1191	D			35	84.2	76.5	CEN	-
1293	S	R	1	1192	S			40	87.1	19.6	CEN	VUS
1294	G			1193	G			35	84.9	17.7	CEN	-
1295	D	E	1	1194	D			40	87.1	18.1	CEN	VUS
1296	E			1195	E			45	71.9	15.1	CEN	-
1297	D	H	1	1196	D			55	93.5	19.6	CEN	VUS
1298	Q			1197	Q			40	92.8	19.6	CEN	-
1299	T	I	1	1198	A			10	A at 90.7	A at 19.0	CEN	LND
1300	Q			1199	Q			40	95	19.6	CEN	-
1301	E			1200	E			40	71.2	14.8	CEN	-
1302	R			1201	R			35	91.4	19.2	CEN	-

1303	V			1202	V			20	I at 62.6	I at 13.0	CEN	-
1304	N			1203	N	H	2	20	50.4	D at 25.8	CEN	ND
1305	K			1204	K			40	92.1	E at 34.3	CEN	-
1306	L			1205	L			40	93.5	52.3	CEN	-
1307	A			1206	A			25	53.2	30.6	CEN	-
1308	K			1207	K			40	79.9	24.1	CEN	-
1309	I			1208	I			40	83.5	25.9	CEN	-
1310	L			1209	L			45	98.6	40.9	CEN	-
1311	K			1210	K			40	82	R at 25.8	CEN	-
1312	E	D	116	1211	E			25	46.8	V at 26.4	CEN	VUS
1313	E			1212	E			20	N at 24.5	Q at 28.8	CEN	-
1314	E			1213	E			35	41	24.1	CEN	-
1315	V	A (M)	5 (2)	1214	V			35	73.4	N at 21.1	CEN	VUS
1316	S			1215	S			15	26.6	K at 54.7	CEN	-
1317	L			1216	S			5	S at 67.6	S at 25.5	CEN	LND
1318	T			1217	S			15	S at 32.4	E at 33.7	CEN	LND
1319	L			1218	L			95	87.8	89.1	CEN	-
1320	C			1219	C			20	R at 41.7	L at 22.3	CEN	-
1321	S			1220	S			20	A at 39.6	A at 31.2	CEN	-
1322	A	V	1	1221	A			50	88.5	R at 32.4	CEN	VUS
1323	G	D	1	1222	G			70	94.2	61	CEN	VUS
1324	V			1223	V			55	87.8	43.8	CEN	-
1325	G			1224	G	C	2	25	41	R at 75.1	CEN	ND
1326	V			1225	V			40	92.1	R at 74.8	CEN	-
1327	I			1226	I			55	71.2	47.5	CEN	-
1328	S			1227	S			45	97.8	T at 68.7	CEN	-
1329	C			1228	C			45	94.2	F at 69.9	CEN	-
1330	I	K	1	1229	I			45	98.6	43.8	CEN	VUS
1331	I			1230	I			40	72.7	28.3	CEN	-
1332	Q			1231	Q			40	97.1	G at 29.7	CEN	-
1333	R			1232	R			55	99.3	29.5	CEN	-
1334	D			1233	D			45	96.4	K at 24.3	CEN	-
1335	E			1234	E			40	82	D at 26.2	CEN	-
1336	G			1235	G			45	71.2	52	CEN	-
1337	R			1236	R			40	82	E at 22.6	CEN	-
1338	T			1237	T			20	M at 27.3	Y at 39.9	CEN	-
1339	P			1238	P			95	97.1	96	CEN	-
1340	M	L	1	1239	M			45	91.4	K at 39.3	CEN	VUS
1341	R			1240	R			45	99.3	Y at 42.7	CEN	-
1342	H			1241	H			45	88.5	F at 47.5	CEN	-
1343	S			1242	S			35	69.1	T at 80.7	CEN	-

1344	F			1243	F			85	95.7	86.2	CEN	-
1345	H			1244	H			25	54	R at 64.3	CEN	-
1346	W			1245	W			40	83.5	G at 32.4	CEN	-
1347	L			1246	S			5	S at 95.0	P at 30.7	CEN	LND
1348	M			1247	L			15	D at 21.6	D at 34.6	CEN	LND
1349	E			1248	E			50	74.1	23.8	CEN	-
1350	K			1249	K			45	95	20.2	CEN	-
1351	Q			1250	Q			15	L at 67.6	L at 14.5	CEN	-
1352	Y			1251	Y			30	68.4	15.1	CEN	-
1353	Y			1252	Y			60	95	58.8	CEN	-
1354	V			1253	V	I	1	20	E at 64.8	E at 39.1	CEN	LND
1355	E	G	116	1254	E			100	100	98.7	CEN	LND
1356	E			1255	E			35	80.6	D at 67.3	CEN	-
1357	P			1256	P			40	87.8	R at 28.6	CEN	-
1358	L			1257	L			20	54.7	I at 36.3	CEN	-
1359	L	Q (M)	11 (3)	1258	L			40	95.7	I at 41.5	CEN	VUS
1360	R			1259	R			100	96.4	98.1	CEN	-
1361	H			1260	H			90	99.3	84.3	CEN	-
1362	V			1261	L			25	L at 53.2	L at 41.4	CEN	LND
1363	E			1262	E			95	100	91.9	CEN	-
1364	P			1263	P			100	97.8	96	CEN	-
1365	P			1264	P			45	99.3	A at 66.6	CEN	-
1366	L			1265	L			90	100	83.8	CEN	-
1367	S			1266	S			45	97.8	A at 73.3	CEN	-
1368	V			1267	I			5	I at 48.2	F at 62.5	CEN	LND
1369	Y			1268	Y			25	48.9	Q at 66.7	CEN	-
1370	L			1269	L			100	98.6	98.5	CEN	-
1371	E			1270	E			100	100	99.4	CEN	-
1372	L			1271	L			100	95.7	94.8	CEN	-
1373	D			1272	D			35	82	G at 30.0	CEN	-
1374	K			1273	K			50	97.8	R at 74.7	CEN	-
1375	L			1274	L			50	84.2	61.8	CEN	-
1376	K	N	1	1275	K			35	80.6	S at 37.0	CEN	VUS
1377	G	V	1	1276	G			20	43.9	N at 50.1	CEN	VUS
1378	Y			1277	Y			35	96.4	F at 61.6	CEN	-
1379	S			1278	S			15	N at 40.3	N at 8.4	CEN	-
1380	N			1279	N			25	48.9	D at 31.5	CEN	-
1381	I			1280	I			25	49.6	46.6	CEN	-
1382	Q	H	2	1281	Q			20	K at 51.1	K at 42.0	CEN	VUS
1383	Y			1282	Y			45	97.8	P at 40.2	CEN	-
1384	T	S	34	1283	T			45	95.7	V at 34.6	CEN	VUS

1385	P			1284	P			85	92.8	58.9	CEN	-
1386	S			1285	S			40	97.1	T at 47.4	CEN	-
1387	R			1286	R			45	98.6	E at 31.0	CEN	-
1388	D			1287	D			50	97.8	N at 68.1	CEN	-
1389	R			1288	R			50	99.3	50.1	CEN	-
1390	Q			1289	Q			45	100	N at 31.8	CEN	-
1391	W			1290	W			45	99.3	I at 37.6	CEN	-
1392	H			1291	H			85	92.1	92.1	CEN	-
1393	M			1292	L			10	L at 50.4	L at 43.5	CEN	LND
1394	Y			1293	Y			95	100	94.5	CEN	-
1395	S			1294	T			5	T at 92.1	E at 31.9	CEN	LND
1396	V			1295	V			60	54	27.9	CEN	-
1397	T			1296	T			20	V at 38.1	A at 34.2	CEN	-
1398	D			1297	D			25	54	K at 66.6	CEN	-
1399	R			1298	K			5	K at 49.6	K at 20.1	CEN	LND
1400	P			1299	P			15	42.5	G at 26.2	CEN	-
1401	V			1300	V	A	1	50	16.6	T at 13.8	CEN	LND
1402	P			1301	P			25	49.6	E at 25.2	CEN	-
1403	I	N	79	1302	I			20	41.7	V at 29.8	CEN	VUS
1404	K			1303	K			35	Q at 45.3	Y at 21.7	CEN	-
1405	R	Q	1	1304	R			100	100	96.1	CEN	PD
1406	M			1305	M			35	68.4	F at 49.5	CEN	-
1407	F			1306	F			100	97.8	97.3	CEN	-
1408	L			1307	L			35	71.9	T at 27.4	CEN	-
1409	R			1308	R			100	100	98.1	CEN	-
1410	S			1309	S			25	T at 85.6	A at 52.6	CEN	-
1411	L			1310	L			30	67.6	I at 48.3	CEN	-
1412	V			1311	V			45	85.6	I at 49.3	CEN	-
1413	R			1312	R			95	99.3	91.8	CEN	-
1414	Q	K	5	1313	Q			45	94.2	P at 31.2	CEN	VUS
1415	T	S	1	1314	A			5	P at 83.5	P at 17.4	CEN	LND
1416	T			1315	T			20	43.2	9.2	CEN	-
1417	M			1316	M			15	A at 25.2	< 5	CEN	-
1418	N			1317	N			25	37.4	8.6	CEN	-
1419	D			1318	D			15	N at 44.6	N at 9.3	CEN	-
1420	G	A	7	1319	G			25	59	12.3	CEN	VUS
1421	F			1320	F			35	77.7	16.3	CEN	-
1422	L			1321	I	M	262	10	T at 36.7	T at 7.7	CEN	ND
1423	L			1322	L			15	S at 49.6	S at 10.3	CEN	-
1424	Q			1323	Q			15	Y at 31.7	< 5	CEN	-
1425	Q			1324	Q			20	59	11.7	CEN	-

1426	G			1325	G			15	I at 38.9	6.3	CEN	-
1427	Q			1326	Q			15	L at 33.1	G at 41.5	CEN	-
1428	D			1327	D	Y	2	30	84.2	52.3	CEN	ND
1429	Y			1328	K			5	V at 22.3	L at 58.2	CEN	LND
1430	Q			1329	Q			15	E at 64.0	R at 32.4	CEN	-
1431	L			1330	L			60	V at 41.7	T at 34.8	CEN	-
1432	S			1331	S			20	G at 34.5	K at 28.6	CEN	-
1433	Q			1332	Q	L	1	15	R at 25.2	I at 32.5	CEN	LND
1434	T			1333	T			25	30.2	< 5	CEN	-
1435	V			1334	L			5	Q at 41.7	A at 29.7	CEN	LND
1436	L			1335	I			10	E at 26.6	S at 29.8	CEN	LND
1437	S			1336	S			70	48.2	A at 31.6	CEN	-
1438	M	I	1	1337	M	I	1	20	L at 40.3	E at 64.9	CEN	LND
1439	A	V	1	1338	A			10	S at 80.6	Y at 67.2	CEN	VUS
1440	F			1339	F			35	82.7	L at 70.0	CEN	-
1441	T			1340	T	M	1	40	91.4	I at 30.1	CEN	LND
1442	S	P	1	1341	S			45	88.5	56.8	CEN	VUS
1443	K	Q	1	1342	K			10	R at 33.8	E at 71.1	CEN	VUS
1444	C	Y	1	1343	C			15	S at 84.9	A at 30.4	CEN	VUS
1445	I			1344	V			35	84.2	17.5	CEN	LND
1446	L			1345	L			35	80.6	E at 33.6	CEN	-
1447	R			1346	R			95	91.4	86.4	CEN	-
1448	S			1347	S			45	94.2	L at 53.4	CEN	-
1449	L			1348	L			80	97.8	57.4	CEN	-
1450	M			1349	M			45	75.5	L at 30.0	CEN	-
1451	N			1350	D			5	A at 35.3	D at 39.6	CEN	LND
1452	A			1351	A			80	99.3	52.6	CEN	-
1453	M			1352	M			60	41	L at 45.4	CEN	-
1454	E			1353	E			50	97.1	D at 69.0	CEN	-
1455	E			1354	E			80	99.3	50.1	CEN	-
1456	L			1355	L			90	84.2	92.4	CEN	-
1457	E			1356	E			95	100	93.6	CEN	-
1458	L			1357	L			45	95	V at 43.9	CEN	-
1459	N			1358	N			25	46	A at 36.0	CEN	-
1460	A			1359	A			40	68.4	F at 29.2	CEN	-
1461	H			1360	H			40	70.5	N at 18.6	CEN	-
1462	N	H	2	1361	N			55	53.2	33.7	CEN	VUS
1463	A			1362	A			15	33.1	T at 34.9	CEN	-
1464	A			1363	A			20	T at 43.2	N at 30.3	CEN	-
1465	M			1364	M			25	I at 41.7	V at 27.1	CEN	-
1466	K	N	9	1365	K			25	48.2	R at 36.0	CEN	VUS

1467	P			1366	P			20	S at 43.9	S at 40.3	CEN	-
1468	D			1367	D			75	43.9	79.2	CEN	-
1469	H			1368	H			45	93.5	C at 33.1	CEN	-
1470	A			1369	A			25	45.3	N at 74.2	CEN	-
1471	H			1370	H			100	95	92.1	CEN	-
1472	M			1371	M			50	96.4	I at 63.9	CEN	-
1473	F			1372	F			80	Y at 79.9	76.9	CEN	-
1474	L			1373	L			90	97.1	57.4	CEN	-
1475	C			1374	C			40	70.5	14.8	CEN	-
1476	I			1375	I			35	84.9	18	CEN	-
1477	L			1376	L			35	75.5	15.9	CEN	-
1478	R	H	1	1377	R			30	63.3	13.3	CEN	VUS
1479	E			1378	E			45	92.1	19.3	CEN	-
1480	Q			1379	Q			45	98.6	20.7	CEN	-
1481	Q			1380	Q	L	1	35	46.8	9.9	CEN	LND
1482	I			1381	I			35	46.8	9.9	CEN	-
1483	D			1382	D			55	31.7	6.8	CEN	-
1484	D			1383	D	A	1	40	95.7	N at 70.3	CEN	LND
1485	L			1384	L			45	95.7	F at 70.2	CEN	-
1486	V			1385	V			30	53.2	44.4	CEN	-
1487	P			1386	P	L	82	45	95	73.6	CEN	ND
1488	Y			1387	F			20	46.8	V at 35.5	CEN	LND
1489	P			1388	P			20	S at 47.5	F at 33.6	CEN	-
1490	R			1389	R			20	K at 34.5	K at 7.2	CEN	-
1491	R			1390	R			25	53.2	T at 28.2	CEN	-
1492	F			1391	V			5	V at 36.0	V at 47.5	CEN	LND
1493	E			1392	E			20	D at 27.3	I at 28.9	CEN	-
1494	V			1393	V	G	1	30	46	M at 26.8	CEN	LND
1495	N			1394	N			20	D at 51.8	D at 39.9	CEN	-
1496	A			1395	A			20	40.3	8.7	CEN	-
1497	E			1396	E			20	G at 65.5	G at 13.6	CEN	-
1498	D			1397	D			15	Q at 69.1	Q at 14.5	CEN	-
1499	E			1398	E			25	51.1	10.8	CEN	-
1500	E	K	1	1399	E			45	95.7	20.1	CEN	VUS
1501	T			1400	T			25	A at 44.6	A at 9.5	CEN	-
1502	T			1401	T			45	68.4	14.7	CEN	-
1503	V			1402	V			20	44.6	Q at 15.9	CEN	-
1504	E			1403	E			25	42.5	P at 62.8	CEN	-
1505	T			1404	M	R	6	15	S at 27.3	S at 26.5	CEN	ND
1506	I			1405	I			25	L at 45.3	K at 27.0	CEN	-
1507	L			1406	L			50	99.3	V at 39.1	CEN	-

1508	E			1407	E			80	51.8	75.4	CEN	-
1509	E			1408	E			70	63.3	55.9	CEN	-
1510	A			1409	A			25	M at 46.0	34.3	CEN	-
1511	T			1410	A			5	A at 78.4	V at 34.3	CEN	LND
1512	Q			1411	R			15	L at 30.9	R at 30.9	CEN	LND
1513	E			1412	E			25	49.6	G at 35.7	CEN	-
1514	I			1413	I			45	94.2	F at 40.0	CEN	-
1515	H			1414	H			35	76.3	L at 35.7	CEN	-
1516	R			1415	R			15	E at 44.6	E at 35.7	CEN	-
1517	S			1416	S			25	39.6	R at 71.4	CEN	-
1518	V			1417	V			40	85.6	Y at 33.3	CEN	-
1519	G			1418	G			95	98.6	88.9	CEN	-
1520	V			1419	V			35	67.6	S at 23.5	CEN	-
1521	R			1420	R			100	87.8	94	CEN	-
1522	M			1421	M			45	98.6	L at 64.5	CEN	-
1523	H	Y	6	1422	H			45	95	W at 60.3	CEN	VUS
1524	A			1423	R			5	R at 64.0	R at 53.4	CEN	LND
1525	L			1424	L			95	100	95.5	CEN	-
1526	G			1425	G			25	46.8	R at 69.4	CEN	-
1527	V			1426	V			95	96.4	89.5	CEN	-
1528	C			1427	C			40	83.5	T at 28.0	CEN	-
1529	E	K	1	1428	E	K	2	20	Q at 45.3	Q at 47.4	CEN	ND
1530	W			1429	W			45	99.3	A at 57.0	CEN	-
1531	E			1430	E			100	100	99.4	CEN	-
1532	V	A	1	1431	V			50	91.4	I at 45.7	CEN	VUS
1533	R			1432	R			35	K at 84.2	K at 49.0	CEN	-
1534	L			1433	L			45	95.7	I at 58.8	CEN	-
1535	W			1434	W			30	64	N at 23.4	CEN	-
1536	L			1435	L			40	60.4	I at 40.9	CEN	-
1537	V	L	2	1436	V			20	D at 32.4	R at 31.8	CEN	VUS
1538	S			1437	S			25	56.1	D at 30.3	CEN	-
1539	S			1438	S			25	D at 43.2	P at 28.6	CEN	-
1540	G	A	3	1439	G			45	94.2	P at 22.8	CEN	VUS
1541	L			1440	L			20	Q at 41.7	T at 47.8	CEN	-
1542	A	T	10	1441	A			50	96.4	G at 57.3	CEN	VUS
1543	N	K	2	1442	C			15	46.8	16	CEN	LND
1544	G			1443	G			45	93.5	P at 30.1	CEN	-
1545	A			1444	A			40	65.5	I at 21.7	CEN	-
1546	W			1445	W			45	99.3	L at 34.2	CEN	-
1547	R			1446	R			100	100	99.1	CEN	-
1548	V			1447	V			55	79.1	40.8	CEN	-

1549	V			1448	V			40	85.6	34	CEN	-
1550	V			1449	V			35	69.8	I at 51.3	CEN	-
1551	A			1450	A			20	T at 66.2	T at 55.5	CEN	-
1552	N			1451	N			95	91.4	89.8	CEN	-
1553	V			1452	V			55	95	35.4	CEN	-
1554	T			1453	T			45	97.8	S at 67.8	CEN	-
1555	G			1454	G			95	80.6	95.1	CEN	-
1556	R			1455	R			20	H at 82.0	Y at 62.1	CEN	-
1557	T			1456	T			45	98.6	V at 32.2	CEN	-
1558	C			1457	C			45	97.8	L at 31.0	CEN	-
1559	T			1458	T			45	79.9	D at 33.6	CEN	-
1560	V			1459	V			45	83.5	43.6	CEN	-
1561	H	L	3	1460	H			20	47.5	E at 33.0	CEN	VUS
1562	I			1461	I			50	89.9	L at 40.6	CEN	-
1563	Y			1462	Y			100	99.3	97.2	CEN	-
1564	R			1463	R			50	98.6	23.8	CEN	-
1565	E			1464	E			100	98.6	97.9	CEN	-
1566	V			1465	V			75	61.2	47.7	CEN	-
1567	E	K	27	1466	E			45	93.5	K at 32.1	CEN	VUS
1568	A			1467	T			10	D at 74.8	D at 42.1	CEN	LND
1569	T	I	2	1468	P			15	56.8	E at 21.3	CEN	LND
1570	G			1469	G			20	E at 27.3	K at 32.2	CEN	-
1571	R			1470	R			15	S at 30.2	T at 27.6	CEN	-
1572	N			1471	N			25	H at 43.2	G at 64.2	CEN	-
1573	S			1472	S	T	21	20	K at 44.6	Q at 17.1	CEN	ND
1574	L			1473	L			45	55.4	I at 26.1	CEN	-
1575	I			1474	I			20	V at 71.2	W at 31.2	CEN	-
1576	Y			1475	Y			45	98.6	F at 62.2	CEN	-
1577	H			1476	H			40	S at 54.0	39.7	CEN	-
1578	S			1477	S			65	54	69.7	CEN	-
1579	I			1478	I			20	25.2	39.3	CEN	-
1580	T			1479	T			40	S at 40.3	G at 22.3	CEN	-
1581	K			1480	K			15	A at 26.6	51	CEN	-
1582	K			1481	K			70	G at 37.4	Q at 28.3	CEN	-
1583	G			1482	G			95	51.1	88.3	CEN	-
1584	P			1483	P			70	96.4	52.6	CEN	-
1585	L			1484	L			75	97.8	52.6	CEN	-
1586	H	L	1	1485	H			95	95	85.6	CEN	VUS
1587	G			1486	E			75	94.2	49.3	-	LND
1588	T			1487	T			20	V at 63.3	M at 26.7	-	-
1589	L			1488	P			35	P at 48.9	P at 54.3	-	LND

1590	I			1489	I			65	L at 48.9	40.3	-	-
1591	N			1490	S			70	71.9	41.4	-	LND
1592	G			1491	D			5	A at 28.1	T at 63.0	-	LND
1593	Q			1492	Q			20	P at 41.0	P at 81.6	-	-
1594	Y	H	6	1493	Y			100	98.6	97.9	-	VUS
1595	K			1494	K			30	Q at 61.2	P at 37.9	-	-
1596	P			1495	P			45	90.7	T at 59.1	-	-
1597	L			1496	L			45	98.6	K at 69.7	-	-
1598	N			1497	G			10	G at 43.9	D at 35.5	-	LND
1599	N			1498	Y			5	V at 52.5	W at 28.5	-	LND
1600	L			1499	L			75	I at 61.2	78.7	-	-
1601	D	N	1	1500	D			45	94.2	Q at 73.0	-	VUS
1602	R			1501	R			25	46	P at 39.9	-	-
1603	K			1502	Q			80	88.5	91	-	LND
1604	R			1503	R			100	99.3	96.9	-	-
1605	L			1504	L			25	56.1	Y at 40.3	-	-
1606	A			1505	A			35	S at 41.7	K at 36.9	-	-
1607	A			1506	A			100	100	97	-	-
1608	R			1507	R			45	97.8	H at 38.7	-	-
1609	R			1508	R			20	K at 45.3	L at 30.0	-	-
1610	S			1509	S			20	N at 64.0	M at 34.9	-	-
1611	N			1510	N			20	30.2	G at 72.1	-	-
1612	T			1511	T			100	99.3	96.9	-	-
1613	T			1512	T			95	100	72.1	-	-
1614	Y			1513	Y			95	95	96.3	-	-
1615	C			1514	C			45	99.3	V at 64.0	-	-
1616	Y			1515	Y			100	100	96.9	-	-
1617	D			1516	D			100	100	99.6	-	-
1618	F			1517	F			85	100	81	-	-
1619	P			1518	P			100	100	93.3	CT-β	-
1620	L			1519	L			45	98.6	E at 57.3	CT-β	-
1621	A			1520	A			45	89.2	L at 46.5	CT-β	-
1622	F			1521	F			100	100	90.7	CT-β	-
1623	E			1522	G			35	87.1	R at 61.3	CT-β	LND
1624	T			1523	T			45	82.7	Q at 65.8	CT-β	-
1625	A	T	1	1524	A			80	89.9	76.9	CT-β	VUS
1626	L			1525	L			70	82	50.1	CT-β	-
1627	E			1526	E			30	42.5	Q at 26.1	CT-β	-
1628	L			1527	L			10	K at 45.3	K at 33.6	CT-β	-
1629	N			1528	L			5	S at 73.4	S at 36.4	CT-β	LND
1630	W			1529	W			100	98.6	99.3	CT-β	-

1631	A			1530	A			15	42.5	K at 18.1	CT-β	-
1632	S			1531	S			45	62.6	K at 30.1	CT-β	-
1633	Q			1532	Q			25	41.7	A at 23.2	CT-β	-
1634	H			1533	H			15	13.7	S at 11.2	CT-β	-
1635	S			1534	P			15	P at 38.1	P at 8.3	CT-β	LND
1636	G			1535	G			35	S at 43.2	10.2	CT-β	-
1637	V			1536	V			40	15.1	H at 13.6	CT-β	-
1638	R			1537	K	E	1	5	V at 23.7	P at 19.9	CT-β	LND
1639	K			1538	K			30	36	S at 18.9	CT-β	-
1640	P			1539	P			40	37.4	L at 35.4	CT-β	-
1641	C			1540	Y			10	K at 30.9	L at 17.1	CT-β	LND
1642	K			1541	K			20	D at 30.9	P at 47.4	CT-β	-
1643	N			1542	D			20	K at 46.0	D at 35.8	CT-β	LND
1644	R			1543	T			5	C at 17.3	C at 25.5	CT-β	LND
1645	L			1544	L			25	Y at 42.5	43.8	CT-β	-
1646	I			1545	I			40	V at 32.4	E at 14.2	CT-β	-
1647	N			1546	N			10	K at 70.5	T at 19.3	CT-β	-
1648	V			1547	V			30	62.6	17.1	CT-β	-
1649	K			1548	K			20	T at 82.7	T at 37.6	CT-β	-
1650	E			1549	E			50	100	99	CT-β	-
1651	L			1550	L			45	100	97.9	CT-β	-
1652	V			1551	V			30	54	70.3	CT-β	-
1653	F			1552	F			45	96.4	L at 61.2	CT-β	-
1654	S			1553	S			25	A at 75.5	D at 67.2	CT-β	-
1655	N			1554	K			10	D at 54.0	D at 11.4	CT-β	LND
1656	T			1555	P	S	1	5	K at 48.2	D at 32.7	CT-β	LND
1657	E	A	2	1556	E			20	23.7	Q at 24.9	CT-β	VUS
1658	G			1557	G			45	97.1	60.1	CT-β	-
1659	S			1558	S	T	1	30	73.4	N at 21.7	CT-β	LND
1660	L			1559	S	L	1	5	W at 85.6	43	CT-β	LND
1661	G	S	2	1560	G			35	88.5	18.7	CT-β	VUS
1662	T			1561	T	I	1	40	90.7	L at 28.8	CT-β	LND
1663	S			1562	S			25	P at 79.9	P at 16.6	CT-β	-
1664	L	V (F)	10 (1)	1563	L			90	75.5	15.9	CT-β	VUS
1665	I			1564	D			30	V at 56.8	V at 43.6	CT-β	LND
1666	P	L	2	1565	L			35	64.8	E at 51.6	CT-β	LND
1667	V			1566	V			50	54	48.1	CT-β	-
1668	E	D	2	1567	E			45	73.4	N at 34.8	CT-β	VUS
1669	R			1568	R			100	93.5	94.6	CT-β	-
1670	P			1569	P			35	55.4	E at 27.9	CT-β	-
1671	A			1570	P	L	2	20	P at 54.7	P at 78.9	CT-β	ND

1672	G			1571	G			100	92.8	96	CT-β	-
1673	L			1572	L			40	73.4	T at 22.3	CT-β	-
1674	N			1573	N			100	100	97.6	CT-β	-
1675	D			1574	D			35	85.6	24.4	CT-β	-
1676	I			1575	F			45	42.5	40.9	CT-β	LND
1677	G	R	1	1576	G			100	99.3	99.7	CT-β	VUS
1678	M			1577	M			90	83.5	96	CT-β	-
1679	V			1578	V			100	94.2	96.1	CT-β	-
1680	A			1579	A			95	97.8	69.3	CT-β	-
1681	W			1580	W			75	100	78.9	CT-β	-
1682	I			1581	C			15	31.7	K at 24.6	CT-β	LND
1683	L			1582	L	V	263	25	M at 58.3	M at 36.0	CT-β	ND
1684	E			1583	D			25	55.4	T at 47.8	CT-β	LND
1685	M			1584	M			40	74.8	L at 28.9	CT-β	-
1686	S			1585	S			40	77.7	K at 42.1	CT-β	-
1687	T			1586	T			90	100	90.9	CT-β	-
1688	P			1587	P			95	100	95.1	CT-β	-
1689	E	G	1	1588	E			100	99.3	97	CT-β	D
1690	F	C	1	1589	F			40	91.4	Y at 74.7	CT-β	VUS
1691	P			1590	P			100	99.3	95.1	CT-β	-
1692	M	V	1	1591	M			15	S at 43.2	E at 22.3	CT-β	VUS
1693	G			1592	G	R	1	100	100	98.1	CT-β	LND
1694	R			1593	R			100	100	99.1	CT-β	-
1695	K			1594	K			30	T at 30.9	R at 23.4	CT-β	-
1696	L			1595	L			15	I at 85.6	I at 44.8	CT-β	-
1697	L			1596	L			20	45.3	I at 64.0	CT-β	-
1698	I			1597	V			15	V at 79.1	V at 66.4	CT-β	LND
1699	V			1598	I	V	4	50	79.9	I at 57.7	CT-β	ND
1700	A			1599	A			60	84.2	66.9	CT-β	-
1701	N			1600	N			100	100	99.9	CT-β	-
1702	D			1601	D			100	100	100	CT-β	-
1703	V			1602	V			35	71.2	I at 75.0	CT-β	-
1704	T			1603	T			100	100	99.3	CT-β	-
1705	F			1604	F			70	95	48.9	CT-β	-
1706	K			1605	K			45	62.6	40.6	CT-β	-
1707	A			1606	A			45	90.7	I at 66.1	CT-β	-
1708	G			1607	G			100	100	100	CT-β	-
1709	S			1608	S			100	100	98.5	CT-β	-
1710	F			1609	F			100	100	99.9	CT-β	-
1711	G			1610	G			100	100	97.5	CT-β	-
1712	P			1611	P			90	93.5	83.1	CT-β	-

1713	R			1612	R			50	92.8	Q at 30.6	CT-β	-
1714	E			1613	E			100	100	100	CT-β	-
1715	D	N	1	1614	D			100	100	99.3	CT-β	VUS
1716	A	V (P)	3 (2)	1615	A	V	1	45	97.8	21.9	CT-β	LND
1717	F			1616	F			60	97.8	57.9	CT-β	-
1718	F			1617	F			90	100	90.9	CT-β	-
1719	L			1618	L			40	27.3	26.2	CT-β	-
1720	A			1619	A			35	87.1	K at 37.3	CT-β	-
1721	V			1620	V			60	97.8	A at 40.6	CT-β	-
1722	T			1621	T			65	100	56.2	CT-β	-
1723	E			1622	E			60	N at 48.9	60.9	CT-β	-
1724	L			1623	L			75	90.7	70.2	CT-β	-
1725	A			1624	A			100	97.8	96.1	CT-β	-
1726	C			1625	C			45	99.3	R at 76.3	CT-β	-
1727	T			1626	A			10	E at 45.3	K at 34.0	CT-β	LND
1728	K			1627	K			20	K at 48.9	L at 33.4	CT-β	-
1729	K			1628	K			50	98.6	G at 67.9	CT-β	-
1730	L			1629	L			35	79.9	I at 65.4	CT-β	-
1731	P			1630	P			100	100	99.7	CT-β	-
1732	L			1631	L			45	99.3	R at 71.4	CT-β	-
1733	I			1632	I			100	98.6	86.4	CT-β	-
1734	Y			1633	Y			100	100	95.7	CT-β	-
1735	L			1634	L			65	100	66.9	CT-β	-
1736	A			1635	A			75	87.1	S at 53.7	CT-β	-
1737	A			1636	A			95	100	90.7	CT-β	-
1738	N			1637	N			90	84.2	96.4	CT-β	-
1739	S	C	12	1638	S			90	86.3	94.3	CT-β	PD
1740	G			1639	G			100	100	99.7	CT-β	-
1741	A			1640	A			100	100	99.7	CT-β	-
1742	R			1641	R			95	100	96.7	CT-β	-
1743	L			1642	L			55	I at 59.7	I at 76.9	CT-β	-
1744	G			1643	G			100	98.6	99.6	CT-β	-
1745	V			1644	V			40	62.6	L at 46.5	CT-β	-
1746	A			1645	A			100	100	93.4	CT-β	-
1747	E			1646	E			80	65.5	68.8	CT-β	-
1748	E			1647	E			100	100	95.5	CT-β	-
1749	V			1648	V			35	81.3	L at 35.2	CT-β	-
1750	K			1649	K			60	95.7	42.6	CT-β	-
1751	A			1650	A			35	S at 51.8	P at 29.4	CT-β	-
1752	C			1651	C			45	97.1	L at 25.0	CT-β	-
1753	F			1652	F			100	100	91.9	CT-β	-

1754	K			1653	K			30	R at 41.0	27.6	CT-β	-
1755	V	D	1	1654	V			90	95	86.4	CT-β	VUS
1756	G			1655	G			45	98.6	A at 69.3	CT-β	-
1757	W			1656	W			100	99.3	94.8	CT-β	-
1758	S	L	193	1657	S	A	1	45	85.6	N at 33.7	CT-β	LND
1759	D			1658	D			95	85.6	78.3	CT-β	-
1760	E			1659	E			30	64	P at 49.3	CT-β	-
1761	V			1660	I			20	S at 40.3	E at 28.3	CT-β	LND
1762	S			1661	S			45	71.9	D at 31.0	CT-β	-
1763	P			1662	P			100	100	88.2	CT-β	-
1764	G			1663	E			5	E at 92.1	E at 43.3	CT-β	LND
1765	N			1664	N			20	R at 71.9	K at 51.7	CT-β	-
1766	G	D (S) (C)	34 (5) (2)	1665	G			100	99.3	97.6	CT-β	LND
1767	F			1666	F			90	99.3	82.9	CT-β	-
1768	Q			1667	Q			50	69.8	K at 47.8	CT-β	-
1769	Y			1668	Y			100	100	97.3	CT-β	-
1770	I			1669	I			40	59	L at 73.8	CT-β	-
1771	Y			1670	Y			100	100	99.3	CT-β	-
1772	L			1671	L			100	97.1	95.4	CT-β	-
1773	S			1672	S	T	1	25	T at 69.8	T at 71.1	CT-β	LND
1774	S			1673	P			10	P at 46.8	P at 58.2	CT-β	LND
1775	E			1674	E			50	82	55.9	CT-β	-
1776	D			1675	D			85	98.6	59.7	CT-β	-
1777	Y			1676	H			65	64	53.1	CT-β	LND
1778	A	T	3	1677	E	K	2	30	54	K at 29.7	CT-β	ND
1779	R			1678	R			70	90.7	48.1	CT-β	-
1780	I			1679	I			50	79.1	L at 32.8	CT-β	-
1781	G			1680	G			25	48.9	S at 30.9	CT-β	-
1782	S			1681	S			55	81.3	39.4	CT-β	-
1783	S			1682	S			50	97.8	21.3	CT-β	-
1784	V			1683	V			45	100	N at 26.5	CT-β	-
1785	I			1684	I			40	90.7	S at 27.0	CT-β	-
1786	A			1685	A			45	98.6	V at 67.8	CT-β	-
1787	H			1686	H			75	99.3	40.8	CT-β	-
1788	E	K	1	1687	E			40	69.1	T at 34.2	CT-β	VUS
1789	V			1688	V			10	L at 59.7	E at 55.0	CT-β	-
1790	K			1689	K			30	63.3	H at 23.4	CT-β	-
1791	L			1690	L			40	85.6	V at 35.2	CT-β	-
1792	P			1691	P	S (L)	386 (1)	20	E at 38.9	E at 43.8	CT-β	ND
1793	S			1692	S			35	66.9	E at 26.5	CT-β	-
1794	G			1693	G	A (T)	80 (1)	95	100	94.8	CT-β	ND

1795	E			1694	E			95	96.4	91.6	CT-β	-
1796	T			1695	T			35	53.2	S at 24.9	CT-β	-
1797	R			1696	R			100	100	94.8	CT-β	-
1798	W			1697	W			45	98.6	Y at 45.4	CT-β	-
1799	V			1698	V			50	83.5	K at 44.7	CT-β	-
1800	I			1699	I			85	82	87.3	CT-β	-
1801	D	N	41	1700	D			45	95.7	T at 46.8	CT-β	VUS
1802	T			1701	T			50	59.7	38.7	CT-β	-
1803	I	F	1	1702	I			85	69.1	81.3	CT-β	VUS
1804	V			1703	V			50	89.2	I at 63.0	CT-β	-
1805	G			1704	G			100	100	99.6	CT-β	-
1806	K			1705	K	E	3	80	92.8	51.9	CT-β	ND
1807	E			1706	E			80	94.2	61.3	CT-β	-
1808	D			1707	D			65	97.1	66.9	CT-β	-
1809	G			1708	G			95	99.3	89.7	CT-β	-
1810	L			1709	I			80	86.3	85.9	CT-β	LND
1811	G			1710	G			100	100	99.4	CT-β	-
1812	V			1711	V			80	84.2	87.3	CT-β	-
1813	E			1712	E			100	100	98.4	CT-β	-
1814	N			1713	N			80	93.5	54.7	CT-β	-
1815	L			1714	L			90	74.8	94.3	CT-β	-
1816	T			1715	T			20	H at 43.2	R at 54.7	CT-β	-
1817	G			1716	G			95	100	89.5	CT-β	-
1818	S			1717	S			95	99.3	91.6	CT-β	-
1819	G			1718	G			90	72.7	93.6	CT-β	-
1820	A			1719	A			45	99.3	L at 49.0	CT-β	-
1821	I	V	1	1720	I			100	100	98.2	CT-β	PD
1822	A			1721	A			100	100	99.6	CT-β	-
1823	G			1722	G			80	S at 60.4	86.4	CT-β	-
1824	A			1723	A			60	96.4	49.9	CT-β	-
1825	Y			1724	Y			45	98.6	T at 57.3	CT-β	-
1826	S			1725	S			100	100	98.1	CT-β	-
1827	R			1726	K			45	79.1	54.3	CT-β	LND
1828	A			1727	A			100	99.3	96.3	CT-β	-
1829	Y			1728	Y			100	100	97.9	CT-β	-
1830	N			1729	N			20	K at 42.5	E at 31.5	CT-β	-
1831	E			1730	E			70	99.3	50.5	CT-β	-
1832	T			1731	T	I	2	45	97.8	I at 67.9	CT-β	ND
1833	F			1732	F			65	100	67.9	CT-β	-
1834	T	S	2	1733	T			100	100	99.4	CT-β	PD
1835	L			1734	L			45	97.1	I at 57.4	CT-β	-

1836	T			1735	T			65	99.3	65.1	CT-β	-
1837	F			1736	F			40	56.1	L at 56.4	CT-β	-
1838	V			1737	V			100	100	97	CT-β	-
1839	S			1738	S			20	T at 83.5	T at 87.6	CT-β	-
1840	G			1739	G			45	98.6	C at 63.7	CT-β	-
1841	R			1740	R			100	97.1	98.7	CT-β	-
1842	S	A	2	1741	T			20	T at 80.6	49.9	CT-β	LND
1843	V			1742	V			60	92.1	71.1	CT-β	-
1844	G			1743	G			100	100	99.7	CT-β	-
1845	I			1744	I			100	100	99.6	CT-β	-
1846	G			1745	G			100	100	99.6	CT-β	-
1847	A			1746	A			90	100	92.4	CT-β	-
1848	Y			1747	Y			100	100	99.6	CT-β	-
1849	L			1748	L			95	100	94.9	CT-β	-
1850	A			1749	A			45	93.5	V at 71.4	CT-β	-
1851	R			1750	R			100	100	99.7	CT-β	-
1852	L			1751	L			100	99.3	98.5	CT-β	-
1853	G			1752	G			100	100	97.3	CT-β	-
1854	M			1753	M			35	71.2	Q at 67.9	CT-β	-
1855	R			1754	R			100	100	99.7	CT-β	-
1856	C			1755	C			45	99.3	A at 35.5	CT-β	-
1857	I			1756	I			90	98.6	82.3	CT-β	-
1858	Q	K	1	1757	Q			100	100	99.6	CT-β	VUS
1859	R			1758	R			45	99.3	V at 46.2	CT-β	-
1860	L			1759	L			40	80.6	E at 63.7	CT-β	-
1861	D			1760	D			50	99.3	G at 39.0	CT-β	-
1862	Q			1761	Q			60	100	58.3	CT-β	-
1863	P			1762	P			65	100	69.7	CT-β	-
1864	I			1763	I			90	99.3	86.1	CT-β	-
1865	I			1764	I			100	100	97.8	CT-β	-
1866	L			1765	L			100	100	99.4	CT-β	-
1867	T			1766	T			100	100	99.6	CT-β	-
1868	G			1767	G			100	100	99.6	CT-β	-
1869	F			1768	F			40	80.6	A at 60.0	CT-β	-
1870	S			1769	S			55	98.6	P at 32.4	CT-β	-
1871	T			1770	T			20	A at 84.2	A at 94.9	CT-β	-
1872	L			1771	L			90	100	73.8	CT-β	-
1873	N			1772	N			100	100	99.4	CT-β	-
1874	K			1773	K			95	100	87.1	CT-β	-
1875	L			1774	L			50	99.3	58.6	CT-β	-
1876	L			1775	L			100	100	96.1	CT-β	-

1877	G			1776	G			100	99.3	99.3	CT-β	-
1878	R			1777	R			95	100	87.3	CT-β	-
1879	E			1778	E			85	100	88.7	CT-β	-
1880	V			1779	V			95	100	96.7	CT-β	-
1881	Y			1780	Y			100	100	99.4	CT-β	-
1882	S			1781	S			45	98.6	T at 58.6	CT-β	-
1883	S	T	8	1782	S			100	100	97	CT-β	PD
1884	H			1783	H			40	88.5	N at 75.7	CT-β	-
1885	M			1784	M			45	97.1	L at 41.1	CT-β	-
1886	Q			1785	Q			100	100	99.4	CT-β	-
1887	L			1786	L			95	100	97.5	CT-β	-
1888	G			1787	G			100	100	99.4	CT-β	D
1889	G			1788	G			100	100	99.4	CT-β	-
1890	P			1789	P			45	97.1	T at 42.1	CT-β	-
1891	K			1790	K			45	99.3	Q at 73.6	CT-β	-
1892	I			1791	I			95	95.7	92.7	CT-β	-
1893	M			1792	M			100	100	99.4	CT-β	-
1894	G			1793	G			20	A at 82.0	Y at 41.5	CT-β	-
1895	T			1794	T			45	93.5	N at 29.1	CT-β	-
1896	N			1795	N			100	100	99.4	CT-β	-
1897	G	S	2	1796	G			100	100	99.4	CT-α	PD
1898	V			1797	V			90	100	92.8	CT-α	-
1899	V			1798	V			45	95.7	S at 54.4	CT-α	-
1900	H			1799	H			100	100	99	CT-α	-
1901	L			1800	L			60	94.2	44.1	CT-α	-
1902	T	K	21	1801	T	A	1	95	100	87.6	CT-α	LD
1903	V			1802	V			75	99.3	47.1	CT-α	-
1904	S			1803	S			45	83.5	25.9	CT-α	-
1905	D			1804	D			90	99.3	77.1	CT-α	-
1906	D			1805	D			95	99.3	95.1	CT-α	-
1907	L			1806	L			55	99.3	46.9	CT-α	-
1908	E			1807	E	A	1	75	96.4	61.3	CT-α	LND
1909	G			1808	G			95	99.3	90.7	CT-α	-
1910	V			1809	V			90	80.6	71.7	CT-α	-
1911	S			1810	S			55	93.5	33	CT-α	-
1912	A			1811	A			30	59.7	K at 32.5	CT-α	-
1913	I			1812	I			85	98.6	80.5	CT-α	-
1914	L			1813	L			80	95	59.1	CT-α	-
1915	N			1814	N			20	K at 51.8	E at 29.7	CT-α	-
1916	W			1815	W			100	99.3	98.4	CT-α	-
1917	L			1816	L			85	99.3	74.5	CT-α	-

1918	S			1817	S			100	99.3	91.8	CT-α	-
1919	Y			1818	Y			90	81.3	69.9	CT-α	-
1920	I			1819	I			30	V at 71.2	V at 53.7	CT-α	-
1921	P			1820	P			100	99.3	99.1	CT-α	-
1922	A			1821	A			45	42.5	28.6	CT-α	-
1923	Y			1822	Y			30	38.1	K at 31.6	CT-α	-
1924	V			1823	V	A (L)	4 (1)	40	41.7	R at 33.0	CT-α	ND
1925	G	S	2	1824	G			50	99.3	47.4	CT-α	VUS
1926	G			1825	G			45	98.6	S at 30.6	CT-α	-
1927	P			1826	P			75	74.8	73.3	CT-α	-
1928	L			1827	L			50	97.1	V at 43.9	CT-α	-
1929	P			1828	P			100	99.3	94.9	CT-α	-
1930	V			1829	V			15	I at 77.0	I at 65.1	CT-α	-
1931	L			1830	L			45	33.1	34.5	CT-α	-
1932	A	V	1	1831	A			15	K at 50.4	S at 16.8	CT-α	VUS
1933	P			1832	P			50	70.5	50.4	CT-α	-
1934	L			1833	L			45	80.6	18	CT-α	-
1935	D			1834	D			100	99.3	98.8	CT-α	-
1936	P			1835	P			75	92.1	55	CT-α	-
1937	P			1836	P			45	88.5	W at 40.3	CT-α	-
1938	E			1837	E			35	59	D at 71.1	CT-α	-
1939	R			1838	R			100	99.3	99.1	CT-α	-
1940	T	S	1	1839	I	T	153	15	P at 68.4	D at 36.4	CT-α	ND
1941	V			1840	V			70	97.1	53.7	CT-α	-
1942	E			1841	E			55	52.5	36.9	CT-α	-
1943	Y			1842	Y	D	2	50	99.3	56.1	CT-α	ND
1944	I			1843	V	I	15	15	F at 37.4	V at 19.9	CT-α	-
1945	P			1844	P			100	98.6	96	CT-α	-
1946	E			1845	E	K	1	45	95.7	P at 30.3	CT-α	LND
1947	N			1846	N			50	88.5	K at 46.2	CT-α	-
1948	S			1847	S			35	66.9	21.1	CT-α	-
1949	C			1848	C			45	98.6	Y at 74.4	CT-α	-
1950	D			1849	D			95	99.3	95.8	CT-α	-
1951	P			1850	P			75	84.2	62.5	CT-α	-
1952	R			1851	R			95	98.6	98.7	CT-α	-
1953	A			1852	A			45	97.8	W at 62.2	CT-α	-
1954	A			1853	A			45	98.6	M at 41.4	CT-α	-
1955	I			1854	I	V	1	65	97.8	60.6	CT-α	LND
1956	A			1855	A			50	S at 36.7	44.4	CT-α	-
1957	G			1856	G			100	99.3	97.9	CT-α	-
1958	I			1857	V	I	1	20	38.9	R at 39.3	CT-α	LND

1959	N			1858	K	N	2	15	D at 28.1	E at 24.7	CT-α	ND
1960	D			1859	D			60	87.8	40.6	CT-α	-
1961	N	D	7	1860	N			20	S at 37.4	P at 28.5	CT-α	VUS
1962	T			1861	T			40	Q at 40.3	19	CT-α	-
1963	G			1862	G			55	98.6	44.9	CT-α	-
1964	K			1863	K			40	76.3	G at 33.9	CT-α	-
1965	W			1864	W			80	99.3	57.1	CT-α	-
1966	L			1865	L			60	85.6	41.7	CT-α	-
1967	G			1866	G			40	89.9	S at 48.9	CT-α	-
1968	G			1867	G			100	99.3	99	CT-α	D
1969	I			1868	I			25	51.1	L at 37.6	CT-α	-
1970	F			1869	F			100	99.3	95.1	CT-α	-
1971	D			1870	D			100	99.3	99	CT-α	-
1972	K			1871	K			55	73.4	50.8	CT-α	-
1973	N			1872	N			20	D at 66.9	G at 48.9	CT-α	-
1974	S			1873	S			100	97.8	97	CT-α	-
1975	F			1874	F			95	99.3	90.1	CT-α	-
1976	V			1875	I	V	194	35	77.7	28	CT-α	ND
1977	E			1876	E			100	99.3	99	CT-α	-
1978	T			1877	T			65	98.6	53.1	CT-α	-
1979	L			1878	L			55	71.9	64.3	CT-α	-
1980	E			1879	E			45	95.7	G at 26.4	CT-α	-
1981	G			1880	G			65	98.6	70.2	CT-α	-
1982	W			1881	W			100	99.3	99	CT-α	-
1983	A			1882	A			100	99.3	96.1	CT-α	-
1984	R			1883	R			25	51.1	K at 36.3	CT-α	-
1985	T			1884	T			80	92.8	76.5	CT-α	-
1986	V			1885	V			100	98.6	97	CT-α	-
1987	V			1886	V			85	85.6	90.6	CT-α	-
1988	T			1887	T			60	97.8	V at 57.3	CT-α	-
1989	G			1888	G			100	98.6	99	CT-α	-
1990	R			1889	R			100	98.6	99.1	CT-α	-
1991	A			1890	A			100	98.6	96.1	CT-α	-
1992	K			1891	K			45	91.4	R at 76.5	CT-α	-
1993	L			1892	L			95	99.3	97.3	CT-α	-
1994	G			1893	G			100	97.8	96.6	CT-α	-
1995	G			1894	G			100	97.8	99	CT-α	-
1996	I			1895	I			90	95	92.1	CT-α	-
1997	P			1896	P			100	99.3	99.1	CT-α	-
1998	I			1897	V			5	97.1	V at 78.7	CT-α	LND
1999	G			1898	G			100	98.6	97.9	CT-α	-

2000	V			1899	V			90	50.4	83.5	CT-α	-
2001	V			1900	V			40	50.4	I at 72.1	CT-α	-
2002	A			1901	A			95	98.6	81.7	CT-α	-
2003	V			1902	V			100	98.6	94.2	CT-α	-
2004	E			1903	E			100	98.6	97	CT-α	-
2005	T			1904	T			100	97.8	89.7	CT-α	-
2006	Q			1905	Q			40	89.2	R at 75.0	CT-α	-
2007	T			1906	T			90	98.6	63.7	CT-α	-
2008	V			1907	V			80	56.1	70	CT-α	-
2009	M			1908	M			40	91.4	E at 65.4	CT-α	-
2010	H			1909	Q			5	Q at 95.0	N at 27.9	CT-α	LND
2011	V			1910	I			15	41.7	31	CT-α	LND
2012	I			1911	I			75	86.3	47.7	CT-α	-
2013	P	L	3	1912	P			95	97.1	98.5	CT-α	PD
2014	A	E	1	1913	A			100	97.8	99	CT-α	PD
2015	D			1914	D			100	97.8	98.8	CT-α	-
2016	P			1915	P			100	97.8	98.8	CT-α	-
2017	G			1916	G			45	97.1	A at 77.4	CT-α	-
2018	Q			1917	Q			45	97.8	N at 70.6	CT-α	-
2019	L			1918	L			80	87.8	46.2	CT-α	-
2020	D	E	1	1919	D			95	97.1	76.9	CT-α	VUS
2021	S			1920	S			100	98.6	98.1	CT-α	-
2022	H			1921	H			35	76.3	E at 28.6	CT-α	-
2023	E			1922	E			65	98.6	68.7	CT-α	-
2024	R			1923	R			45	94.2	K at 27.1	CT-α	-
2025	V			1924	V			45	65.5	33.6	CT-α	-
2026	V			1925	V			50	97.1	I at 32.7	CT-α	-
2027	P	R	2	1926	P			45	97.1	Q at 37.2	CT-α	VUS
2028	Q			1927	Q			70	70.5	49.6	CT-α	-
2029	A			1928	A			95	98.6	96	CT-α	-
2030	G			1929	G			100	98.6	98.8	CT-α	-
2031	Q			1930	Q			95	98.6	75	CT-α	-
2032	V			1931	V			100	98.6	99	CT-α	-
2033	W			1932	W			100	98.6	98.5	CT-α	-
2034	F			1933	F			75	97.1	52.2	CT-α	-
2035	P			1934	P			100	98.6	99	CT-α	-
2036	D			1935	D			75	95	56.8	CT-α	-
2037	S			1936	S			100	98.6	99	CT-α	-
2038	A			1937	A			95	98.6	92.7	CT-α	-
2039	A			1938	A			25	T at 70.5	F at 52.2	CT-α	-
2040	K			1939	K			100	98.6	99	CT-α	-

2041	T			1940	T			100	98.6	99	CT-α	-
2042	A			1941	A			90	93.5	88.9	CT-α	-
2043	Q			1942	Q			100	98.6	91.9	CT-α	-
2044	A			1943	A			90	98.6	94.3	CT-α	-
2045	L	F	8	1944	L			35	61.9	I at 70.0	CT-α	VUS
2046	M			1945	M			20	L at 60.4	K at 28.3	CT-α	-
2047	D			1946	D			100	98.6	98.8	CT-α	-
2048	F			1947	F			100	97.1	94.5	CT-α	-
2049	N			1948	N			95	98.6	84.9	CT-α	-
2050	R			1949	R			70	89.2	49.9	CT-α	-
2051	E			1950	E			100	97.8	98.8	CT-α	-
2052	Q			1951	E			25	E at 56.1	36	CT-α	LND
2053	L			1952	L			100	98.6	98.7	CT-α	-
2054	P			1953	P			100	98.6	98.7	CT-α	-
2055	L			1954	L			95	98.6	84.6	CT-α	-
2056	F			1955	F			50	97.8	M at 50.5	CT-α	-
2057	I			1956	I			75	98.6	80.5	CT-α	-
2058	I			1957	L			5	L at 95.7	L at 59.1	CT-α	LND
2059	A	V	1	1958	A			100	98.6	98.2	CT-α	VUS
2060	N			1959	N			100	97.8	97.9	CT-α	-
2061	W			1960	W			100	98.6	97.3	CT-α	-
2062	R			1961	R			100	98.6	99	CT-α	-
2063	G			1962	G			100	98.6	99	CT-α	-
2064	F			1963	F			100	98.6	99	CT-α	-
2065	S			1964	S			100	98.6	99	CT-α	-
2066	G			1965	G			100	98.6	98.2	CT-α	-
2067	G			1966	G			100	98.6	99	CT-α	-
2068	Q	R	1	1967	Q			65	98.6	64.9	CT-α	VUS
2069	R			1968	R			65	96.4	62.5	CT-α	-
2070	D			1969	D			100	98.6	99	CT-α	-
2071	L			1970	L			45	97.8	M at 78.3	CT-α	-
2072	F			1971	F			55	97.8	Y at 64.8	CT-α	-
2073	E			1972	E			55	97.8	N at 34.9	CT-α	-
2074	G			1973	G			45	97.8	E at 49.6	CT-α	-
2075	I			1974	I			50	93.5	V at 61.8	CT-α	-
2076	L			1975	L			95	97.8	92.8	CT-α	-
2077	Q	L	9	1976	Q			45	97.1	K at 78.3	CT-α	VUS
2078	A			1977	A			45	97.8	F at 37.0	CT-α	-
2079	G			1978	G			100	98.6	99	CT-α	-
2080	S			1979	S			70	93.5	62.8	CT-α	-
2081	A			1980	T			10	T at 79.9	Y at 44.5	CT-α	LND

2082	I	M	1	1981	I			95	97.1	98.7	CT-α	VUS
2083	V	G	1	1982	V			100	98.6	98.8	CT-α	VUS
2084	E			1983	E			45	97.8	D at 78.1	CT-α	-
2085	N			1984	N			45	93.5	A at 40.9	CT-α	-
2086	L			1985	L			100	98.6	97.8	CT-α	-
2087	R			1986	R			80	97.1	53.1	CT-α	-
2088	T			1987	T			45	94.2	K at 25.2	CT-α	-
2089	Y			1988	Y			80	97.8	82.9	CT-α	-
2090	R			1989	R			20	K at 43.9	K at 40.0	CT-α	-
2091	Q			1990	Q			95	98.6	88.2	CT-α	-
2092	P			1991	P			100	98.6	99	CT-α	-
2093	V			1992	V			55	43.2	52.3	CT-α	-
2094	F			1993	F			55	98.6	61.5	CT-α	-
2095	V	L	1	1994	V			70	95.7	67.9	CT-α	VUS
2096	Y			1995	Y			100	92.8	96.1	CT-α	-
2097	I			1996	I			95	97.1	92.1	CT-α	-
2098	P	S	2	1997	P			100	97.8	93	CT-α	VUS
2099	M			1998	M			25	64.8	P at 77.4	CT-α	-
2100	M	I	2	1999	M			25	48.9	F at 21.3	CT-α	VUS
2101	G			2000	G			55	79.9	65.4	CT-α	-
2102	E			2001	E			100	98.6	98.8	CT-α	-
2103	L			2002	L			100	98.6	98.2	CT-α	-
2104	R			2003	R			100	98.6	99	CT-α	-
2105	G			2004	G			100	98.6	99	CT-α	-
2106	G			2005	G			100	98.6	99	CT-α	-
2107	A			2006	A			50	98.6	S at 56.2	CT-α	-
2108	W			2007	W			100	98.6	99	CT-α	-
2109	V			2008	V			95	94.2	87.7	CT-α	-
2110	V			2009	V			100	98.6	98.8	CT-α	-
2111	V			2010	V			60	76.3	54.9	CT-α	-
2112	D			2011	D			100	97.8	98.5	CT-α	-
2113	S	T	11	2012	S			60	97.8	P at 62.2	CT-α	VUS
2114	Q			2013	Q			20	K at 60.4	T at 59.1	CT-α	-
2115	I	R	1	2014	I			100	97.8	98.2	CT-α	PD
2116	N			2015	N			100	98.6	98.1	CT-α	-
2117	S	L	1	2016	S			25	P at 57.6	P at 62.1	CT-α	VUS
2118	D	E	2	2017	D			50	83.5	35.8	CT-α	VUS
2119	Y			2018	Y			25	H at 53.2	H at 27.0	CT-α	-
2120	I			2019	V			40	80.6	M at 67.9	CT-α	LND
2121	E			2020	E			100	98.6	96.7	CT-α	-
2122	M	I	4	2021	M			80	70.5	81.9	CT-α	VUS

2123	Y			2022	Y			100	98.6	93.1	CT-α	-
2124	A			2023	A			100	97.8	94.3	CT-α	-
2125	D			2024	D			75	E at 64.8	84.1	CT-α	-
2126	E	K	51	2025	E			25	R at 75.6	R at 26.5	CT-α	VUS
2127	T	A	1	2026	T			45	96.4	E at 45.1	CT-α	VUS
2128	A			2027	A			55	95	S at 24.9	CT-α	-
2129	R			2028	R	L (H)	5 (1)	80	K at 67.6	84	CT-α	ND
2130	G			2029	G			90	96.4	80.1	CT-α	-
2131	N			2030	N			45	95	G at 70.6	CT-α	-
2132	V			2031	V			100	97.1	86.5	CT-α	-
2133	L			2032	L			100	98.6	97.6	CT-α	-
2134	E			2033	E			100	98.6	98.5	CT-α	-
2135	P			2034	P			90	87.1	89.8	CT-α	-
2136	E			2035	E			75	60.4	71.8	CT-α	-
2137	G			2036	G			100	97.8	95.8	CT-α	-
2138	M			2037	T			25	L at 46.0	I at 34.3	CT-α	LND
2139	I			2038	I			45	59.9	V at 65.1	CT-α	-
2140	E			2039	E			80	98.6	63.1	CT-α	-
2141	I			2040	I			100	98.6	93.6	CT-α	-
2142	K			2041	K			100	97.1	96.7	CT-α	-
2143	F			2042	F			85	97.1	53.2	CT-α	-
2144	R			2043	R			90	88.5	90.9	CT-α	-
2145	R			2044	T			20	T at 38.1	33.3	CT-α	LND
2146	K	Q	15	2045	K			55	41	39	CT-α	VUS
2147	E			2046	E			45	89.2	K at 36.3	CT-α	-
2148	L			2047	L			80	88.5	59.5	CT-α	-
2149	L			2048	L			45	49.6	51.4	CT-α	-
2150	E			2049	E			45	79.1	K at 31.2	CT-α	-
2151	C			2050	C			40	80.6	T at 62.5	CT-α	-
2152	M	L	1	2051	M			90	98.6	89.2	CT-α	VUS
2153	G			2052	G			30	74.8	A at 24.1	CT-α	-
2154	R			2053	R			100	97.8	97.6	CT-α	-
2155	L			2054	L			65	95.7	65.8	CT-α	-
2156	D			2055	D			100	98.6	96.9	CT-α	-
2157	Q			2056	Q			20	P at 50.4	P at 55.0	CT-α	-
2158	T			2057	K			15	E at 43.2	V at 23.7	CT-α	LND
2159	L			2058	L			45	96.4	Y at 60.0	CT-α	-
2160	I			2059	I	V	1	60	90.7	29.7	CT-α	LND
2161	N			2060	S			15	38.9	E at 17.7	CT-α	LND
2162	L			2061	L			100	94.2	92.5	CT-α	-
2163	K			2062	K			45	76.3	44.1	CT-α	-

2164	A			2063	A			35	71.9	E at 26.4	CT-α	-
2165	N			2064	K			5	K at 70.5	K at 26.1	CT-α	LND
2166	I			2065	L			5	L at 93.5	L at 76.9	CT-α	LND
2167	Q	E	3	2066	Q	R	2	20	65.5	D at 27.1	CT-α	ND
2168	D			2067	D			30	E at 43.2	E at 12.1	CT-α	-
2169	A			2068	A			45	67.6	A at 21.6	CT-α	-
2170	K			2069	K			40	58.3	G at 18	CT-α	-
2171	R	Q (G)	2 (1)	2070	Q			5	S at 30.2	T at 15.1	CT-α	LND
2172	N			2071	S			10	S at 21.6	S at 30.4	CT-α	LND
2173	K			2072	E	V	1	15	N at 23.7	L at 47.8	CT-α	LND
2174	A	S	18	2073	A			25	G at 25.9	S at 50.1	CT-α	VUS
2175	Y			2074	Y			15	S at 36.7	P at 16.2	CT-α	-
2176	A			2075	A			20	31.7	E at 35.1	CT-α	-
2177	N			2076	N			15	D at 23.0	E at 37.5	CT-α	-
2178	I			2077	I			20	V at 20.1	R at 21.9	CT-α	-
2179	E			2078	E			40	69.8	K at 23.1	CT-α	-
2180	L			2079	L			20	S at 51.1	E at 36.1	CT-α	-
2181	L			2080	L			60	69.1	45	CT-α	-
2182	Q			2081	Q			25	72.7	K at 27.6	CT-α	-
2183	K			2082	Q			25	Q at 41.0	23.5	CT-α	LND
2184	Q	H	1	2083	Q			35	48.2	K at 40.5	CT-α	VUS
2185	I			2084	I	V	262	40	82.7	L at 42.9	CT-α	ND
2186	K			2085	K			60	39.6	37.3	CT-α	-
2187	T	I	1	2086	A			5	A at 56.1	A at 36.6	CT-α	LND
2188	R			2087	R			100	98.6	96.4	CT-α	-
2189	E			2088	E			70	51.8	82.8	CT-α	-
2190	K			2089	K			45	89.9	37	CT-α	-
2191	Q			2090	Q			50	86.3	39.1	CT-α	-
2192	L			2091	L			100	97.8	96.4	CT-α	-
2193	L			2092	L			65	75.5	64.6	CT-α	-
2194	P			2093	P			95	98.6	92.2	CT-α	-
2195	V	I	1	2094	V			25	40.3	I at 42.3	CT-α	VUS
2196	Y			2095	Y			100	98.6	96	CT-α	-
2197	T			2096	I			30	83.5	H at 27.7	CT-α	LND
2198	Q			2097	Q			95	98.6	87.1	CT-α	-
2199	I			2098	I			55	79.1	54.3	CT-α	-
2200	A			2099	A			80	98.6	80.2	CT-α	-
2201	T			2100	T			35	66.2	V at 39.6	CT-α	-
2202	K	T	1	2101	K			25	R at 52.5	Q at 52.9	CT-α	VUS
2203	F			2102	F			100	97.8	91.5	CT-α	-
2204	A			2103	A			95	94.2	92.5	CT-α	-

2205	E			2104	E			40	92.1	D at 76.0	CT-α	-
2206	L			2105	L			95	97.1	93.9	CT-α	-
2207	H	Q	1	2106	H			100	98.6	98.1	CT-α	PD
2208	D			2107	D			100	98.6	98.4	CT-α	-
2209	T			2108	T			80	92.1	50.7	CT-α	-
2210	S			2109	S			55	93.5	P at 32.2	CT-α	-
2211	M			2110	M			20	L at 62.6	G at 59.4	CT-α	-
2212	R			2111	R			95	97.1	94.2	-	-
2213	M			2112	M			100	97.1	97.2	-	-
2214	A			2113	A	E	4	40	89.2	L at 20.5	-	ND
2215	A			2114	A			60	95	65.5	-	-
2216	K			2115	K			100	96.4	90.1	-	-
2217	G			2116	G			90	97.8	85.8	-	-
2218	V			2117	V			90	95.7	65.2	-	-
2219	I			2118	I			95	89.9	84.7	-	-
2220	K			2119	K			30	44.6	R at 46.5	-	-
2221	S	R	22	2120	S			20	K at 48.2	D at 29.5	-	VUS
2222	V			2121	V			55	95.7	33.7	-	-
2223	V			2122	V			45	80.6	L at 56.4	-	-
2224	E			2123	E			50	D at 79.1	31.3	-	-
2225	W			2124	W			100	97.8	96.3	-	-
2226	S	R	1	2125	S			25	E at 49.6	K at 36.0	-	VUS
2227	G			2126	G	S	46	20	E at 31.7	N at 33.6	-	ND
2228	S	L	1	2127	S			70	97.1	53.5	-	VUS
2229	R			2128	R			100	98.6	97.3	-	-
2230	S	L	49	2129	S	A	1	40	65.5	R at 45.7	-	LND
2231	F			2130	F			90	85.6	73	-	-
2232	F			2131	F			85	98.6	82.2	-	-
2233	Y			2132	Y	H	230	95	91.4	82.6	-	ND
2234	K			2133	K			30	46	W at 71.1	-	-
2235	K			2134	K			20	R at 87.1	R at 87.9	-	-
2236	L			2135	L			95	98.6	85.8	-	-
2237	Y			2136	N			5	R at 48.9	R at 79.2	-	LND
2238	R			2137	R			100	98.6	93.3	-	-
2239	R			2138	R			65	97.8	64.9	-	-
2240	I			2139	I			35	60.4	L at 73.0	-	-
2241	A			2140	A			35	44.6	L at 27.6	-	-
2242	E			2141	E			100	95.7	88.8	-	-
2243	S			2142	S			20	D at 46.0	E at 42.0	-	-
2244	S			2143	S			30	56.1	Y at 23.2	-	-
2245	L			2144	L			50	94.2	35.2	-	-

2246	V			2145	V			20	A at 43.2	L at 25.3	-	-
2247	R			2146	K	R	13	15	K at 77.0	K at 54.1	-	ND
2248	N			2147	N			20	E at 33.1	R at 27.7	-	-
2249	I	T	1	2148	V	I	6	60	V at 56.8	40.5	-	ND
2250	R			2149	R			35	79.1	17.5	-	-
2251	K			2150	E			10	D at 30.9	E at 18.1	-	LND
2252	A			2151	A			75	70.5	43	-	-
2253	S			2152	S			35	A at 66.9	A at 16.6	-	-
2254	G			2153	G			55	97.8	23.5	-	-
2255	D			2154	D			30	44.6	9.8	-	-
2256	I			2155	N	S	247	5	Q at 54.0	Q at 11.4	-	ND
2257	L			2156	L			25	54	12	-	-
2258	S			2157	A	T	238	35	61.2	35.7	-	ND
2259	Y			2158	Y			25	H at 69.8	R at 27.1	-	-
2260	K			2159	K			20	46	G at 17.8	-	-
2261	S			2160	S			45	85.6	Q at 20.2	-	-
2262	A			2161	S	A	184	40	97.1	28.2	-	ND
2263	M			2162	M			25	43.2	L at 28.5	-	-
2264	G	V	6	2163	R	G	190	15	E at 46.0	A at 25.8	-	ND
2265	L			2164	L			35	67.6	M at 30.6	-	-
2266	I			2165	I			40	78.4	L at 64.8	-	-
2267	Q			2166	Q			20	K at 77.7	R at 37.2	-	-
2268	D			2167	D			25	K at 43.2	R at 26.8	-	-
2269	W			2168	W			45	95.7	84.1	-	-
2270	F			2169	F			20	Y at 53.2	37.3	-	-
2271	R	C	1	2170	C	S	1	10	L at 74.1	V at 22.2	-	LND
2272	K			2171	N			15	A at 42.5	A at 8.3	-	LND
2273	S			2172	S			45	49.6	19.5	-	-
2274	E			2173	D	V	220	10	S at 47.5	29.8	-	ND
2275	I			2174	I			20	30.9	V at 9.9	-	-
2276	A			2175	A	T	2	25	54	E at 19.6	-	ND
2277	K			2176	K			30	28.1	G at 27.3	-	-
2278	G			2177	G			40	46	A at 15.9	-	-
2279	K			2178	K			35	26.6	23.5	-	-
2280	E			2179	E			30	36	A at 21.9	-	-
2281	E			2180	E			30	D at 32.4	Y at 25.2	-	-
2282	A			2181	A			30	43.9	L at 20.4	-	-
2283	W			2182	W			45	97.8	57.1	-	-
2284	T	R	61	2183	T			20	D at 28.8	D at 35.5	-	VUS
2285	D			2184	D			65	90.7	42.7	-	-
2286	D			2185	D			45	97.8	66	-	-

2287	Q			2186	Q			35	E at 47.5	R at 22.6	-	-
2288	L			2187	V			5	A at 72.7	A at 34.0	-	LND
2289	F			2188	F			45	96.4	V at 53.1	-	-
2290	F			2189	F			35	66.9	A at 32.5	-	-
2291	T			2190	T			25	A at 49.6	E at 27.0	-	-
2292	W			2191	W			45	98.6	85	-	-
2293	K			2192	K			35	73.4	E at 65.1	-	-
2294	D			2193	D			35	58.3	E at 32.4	-	-
2295	N			2194	N			40	D at 45.3	35.5	-	-
2296	V			2195	V	A	6	20	P at 63.3	L at 16.0	-	ND
2297	S			2196	S			55	E at 27.3	K at 25.5	-	-
2298	N			2197	N			35	84.9	22.6	-	-
2299	Y			2198	Y			45	92.1	I at 36.7	-	-
2300	E			2199	E			40	56.8	38.5	-	-
2301	Q			2200	L			5	E at 29.5	E at 26.4	-	LND
2302	K			2201	K			25	Y at 39.6	N at 26.1	-	-
2303	L	V	1	2202	L			40	68.4	I at 27.7	-	VUS
2304	S			2203	S			20	K at 34.5	K at 25.0	-	-
2305	E			2204	E			50	88.5	27.9	-	-
2306	L			2205	L			70	98.6	63	-	-
2307	R			2206	R			45	74.8	K at 48.6	-	-
2308	T			2207	A			5	A at 61.9	R at 23.5	-	LND
2309	Q			2208	Q			35	77.7	D at 36.6	-	-
2310	K			2209	K			25	61.2	14.4	-	-
2311	L			2210	L			30	V at 70.5	V at 37.3	-	-
2312	L			2211	L			55	50.4	31.5	-	-
2313	N			2212	N			20	L at 32.4	K at 22.0	-	-
2314	Q			2213	Q			40	49.6	29.8	-	-
2315	L			2214	L			55	90.7	I at 34.9	-	-
2316	A			2215	A			25	S at 56.1	28.6	-	-
2317	E			2216	E			20	N at 25.2	S at 25.8	-	-
2318	I			2217	I			25	42.5	L at 52.9	-	-
2319	G			2218	G			25	54	V at 26.1	-	-
2320	N			2219	N			20	D at 41.7	Q at 21.4	-	-
2321	S	T	1	2220	S			50	80.6	G at 34.3	-	VUS
2322	S			2221	S			40	64	V at 21.1	-	-
2323	D			2222	D			45	96.4	E at 34.8	-	-
2324	L			2223	L			35	79.1	V at 45.3	-	-
2325	Q	K	1	2224	Q			25	51.8	10.9	-	VUS
2326	A			2225	A			50	92.1	41.2	-	-
2327	L			2226	L			40	87.1	24.4	-	-

2328	P			2227	P			40	86.3	D at 24.6	-	-
2329	Q			2228	Q			35	78.4	16.6	-	-
2330	G			2229	G			55	93.5	19.5	-	-
2331	L			2230	L			55	94.2	21.9	-	-
2332	A			2231	A			35	48.2	V at 15.0	-	-
2333	N			2232	N			20	A at 36.0	H at 13.9	-	-
2334	L			2233	L			55	90.7	56.8	-	-
2335	L			2234	L			60	95.7	21.6	-	-
2336	N			2235	N			20	S at 34.5	Q at 25.8	-	-
2337	K			2236	K	M	32	45	94.2	22	-	ND
2338	V			2237	V			20	M at 49.6	L at 42.4	-	-
2339	D			2238	E			25	45.3	S at 31.5	-	LND
2340	L			2239	P	R	1	5	P at 77.0	P at 44.7	-	LND
2341	S			2240	S			40	68.4	E at 25.2	-	-
2342	R			2241	K			15	K at 26.6	E at 34.2	-	LND
2343	R			2242	R			90	96.4	64.6	-	-
2344	E			2243	E			40	34.5	A at 34.3	-	-
2345	E			2244	E			35	Q at 54.7	Q at 24.7	-	-
2346	L			2245	L			35	52.5	V at 37.2	-	-
2347	V			2246	V			55	41	L at 33.1	-	-
2348	D	N	1	2247	A			20	37.4	K at 28.5	-	LND
2349	A			2248	A			20	E at 61.9	Y at 15.6	-	-
2350	I			2249	I			10	L at 58.3	L at 78.3	-	-
2351	R			2250	R			35	61.9	S at 24.7	-	-
2352	K			2251	K			35	71.2	T at 21.0	-	-
2353	V			2252	V			40	75.5	18.3	-	-
2354	L			2253	L			45	92.8	22.3	-	-
2355	G	S	1	2254	G			30	54	S at 16.3	-	VUS
2356	X			2255	X			-	-	-	-	-

APPENDIX J: ACC2 Consensus Protein Sequence with Conserved Residues and Genetic Variants Highlighted

This appendix shows the consensus ACC2 protein sequence among 857 Arabidopsis accessions. Also shown in this appendix is the conservation percentage of each amino acid based on our multi-kingdom alignment of 667 eukaryotic ACCase sequences, accession variation for ACC1 and ACC2, and the current classification of each variant based upon its likely impact on ACCase function. Adapted from Parker et al. (2016).

Footnotes for each row are described below:

First row: Red, ≥ 99% conserved in the multi-kingdom alignment of 667 eukaryotic ACCase sequences; Purple, 95-98%; Blue, 90-94%; Green, 80-89%; Black < 80%.

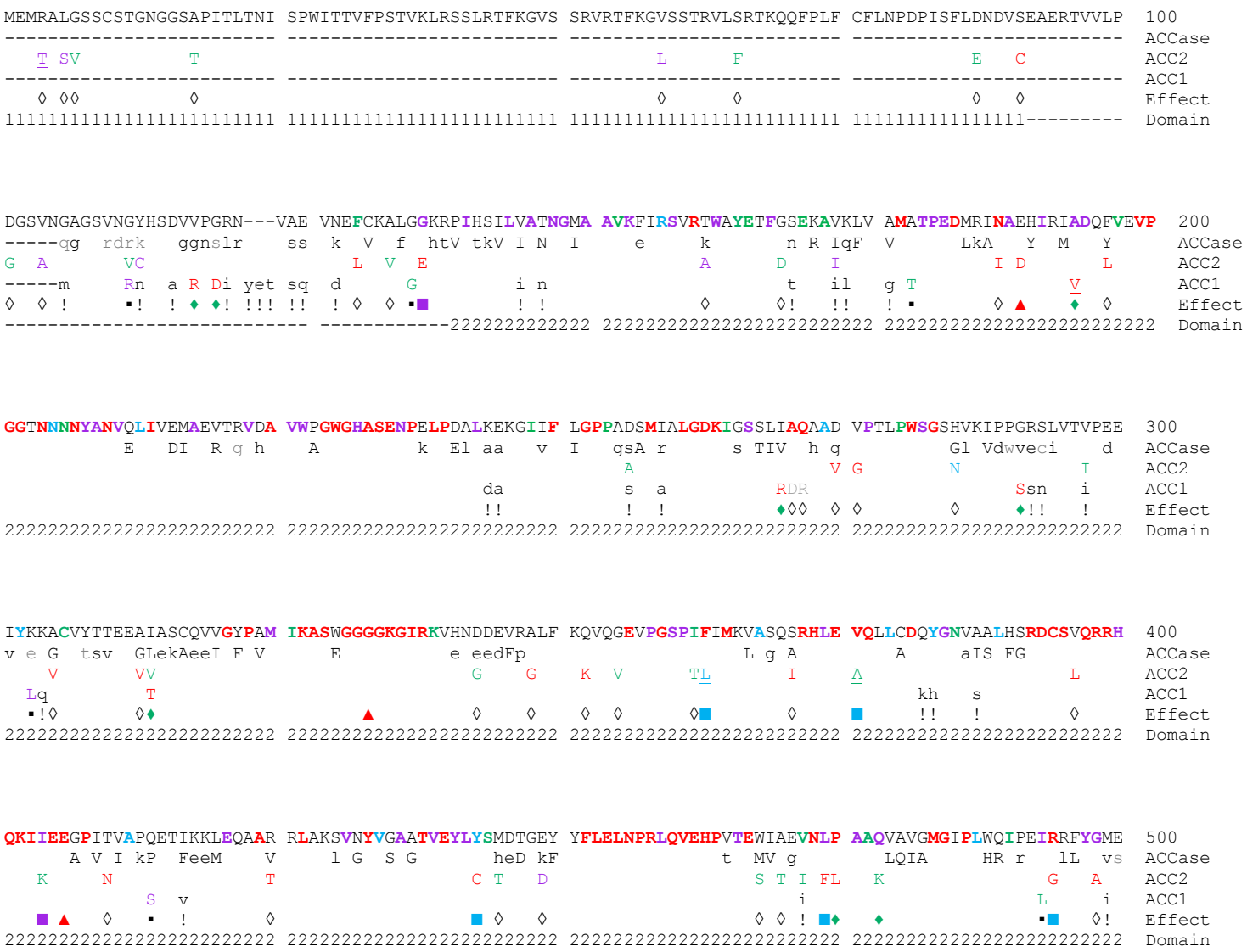
Second row: Residues (consensus from homomeric ACCase alignment) that differ from the ACC2 consensus among sequenced accessions. Capital letters, amino acid indicated is ≥ 50% conserved in the multi-kingdom alignment of 667 eukaryotic ACCase sequences; Lower case letters, < 50%; Gray letters, < 25%. Residues preceding the start of ACC1 are excluded.

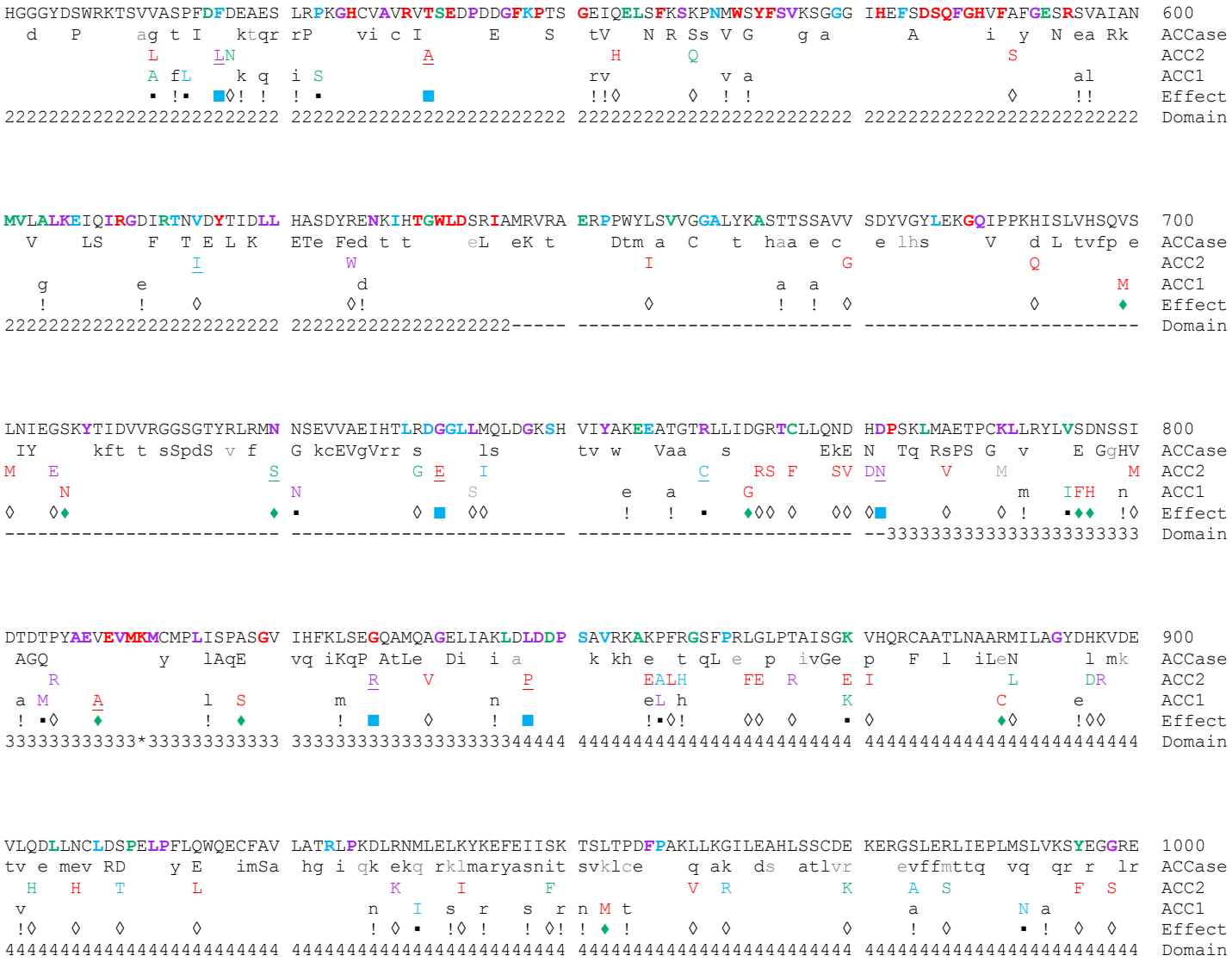
Third row: Most common ACC2 variant identified among sequenced accessions; Red, 1 accession with variant indicated; Purple, 2-3; Blue, 4-10; Green, >10; Gray, variant not found in the single accession predicted; Underlined, variant confirmed by Sanger sequencing.

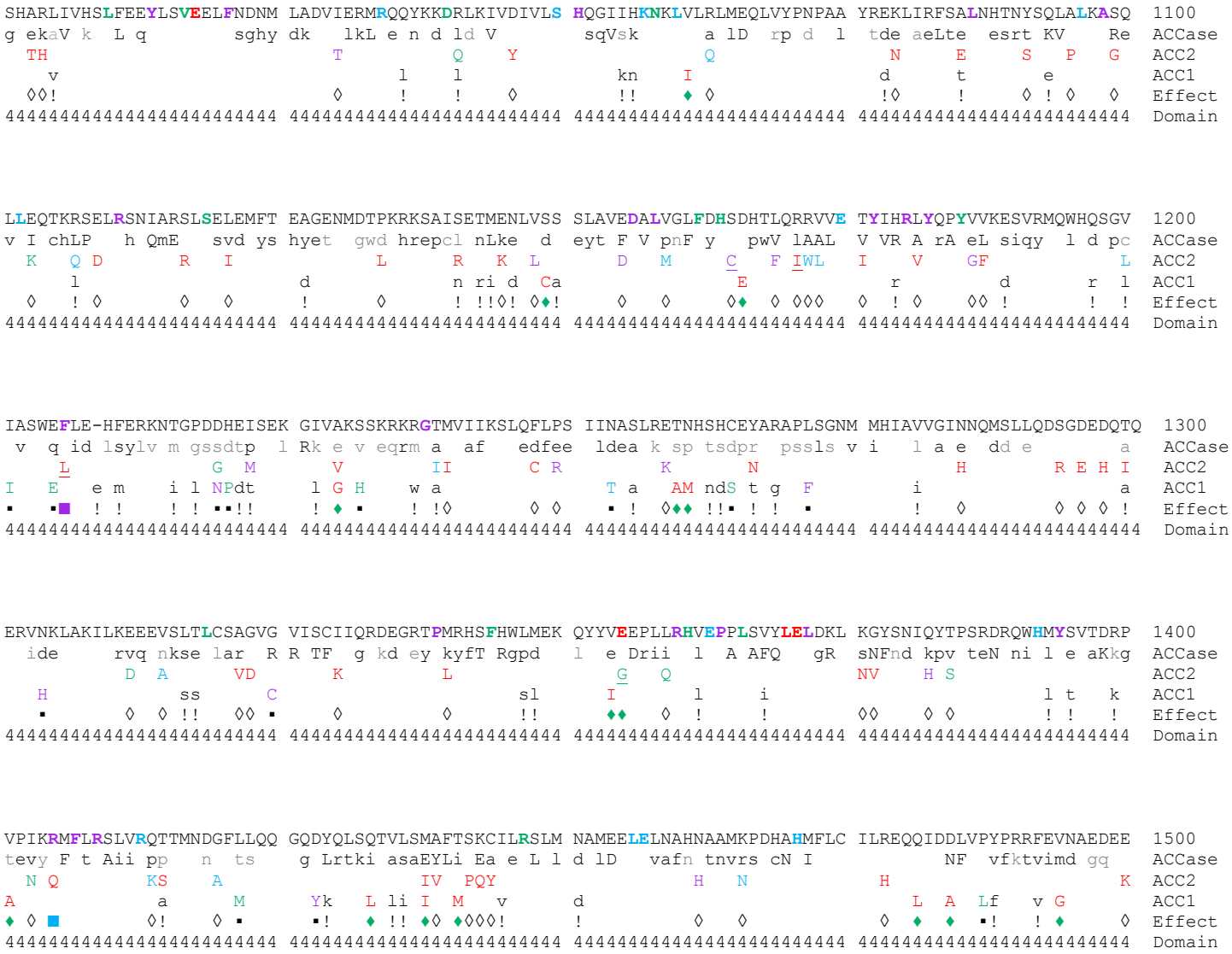
Fourth row: Most common ACC1 variant identified; colors and underlining same as for ACC2 above. Lower case letters, consensus ACC1 residue differs from that found in ACC2.

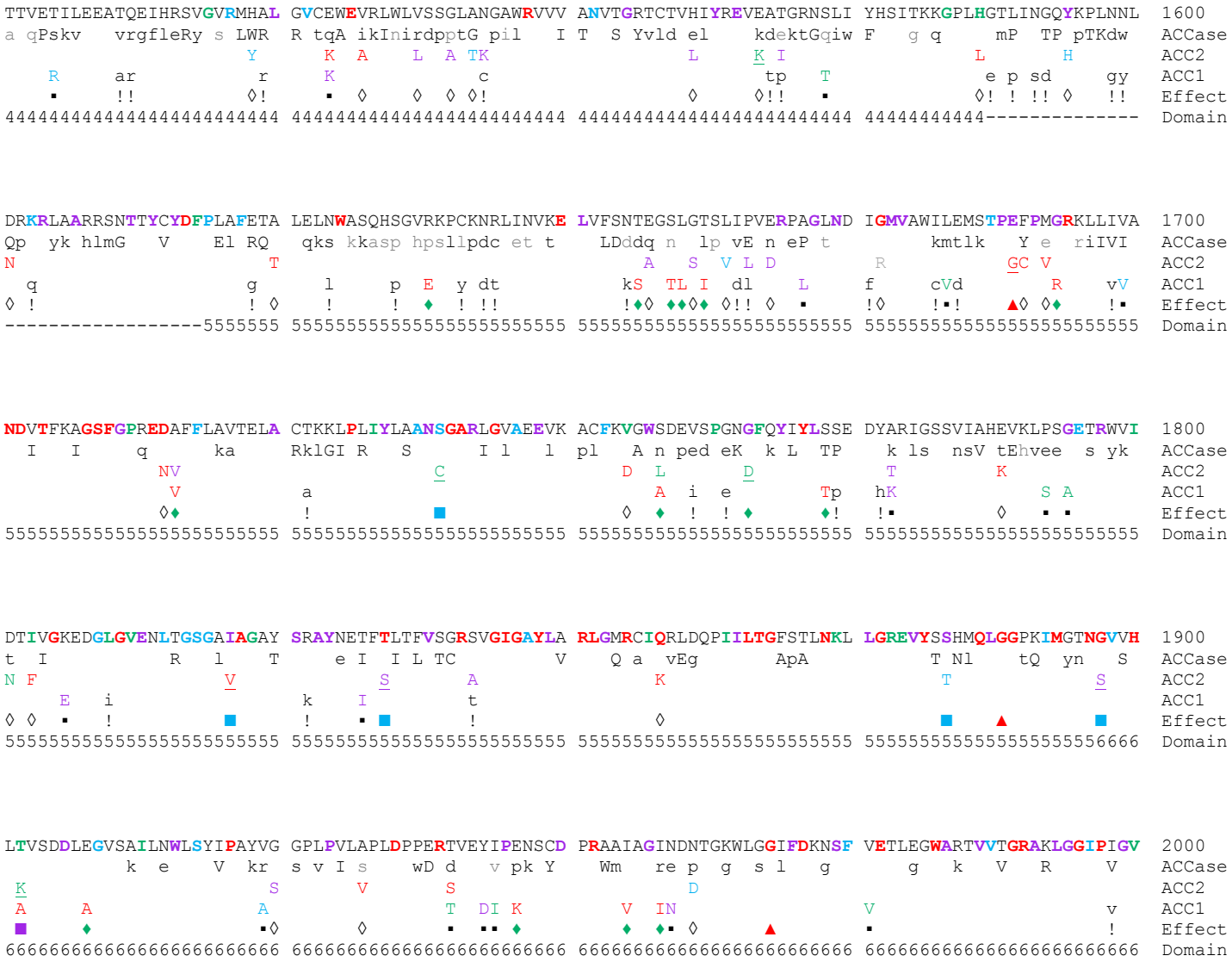
Fifth row: Variant classification. Red square, deleterious to protein function; Red triangle, deleterious based on phenotype of induced acc1 missense mutation; Purple square, likely deleterious; Blue square, possibly deleterious; Open diamond, variant of unknown significance; Exclamation point, Likely not deleterious (ACC1 consensus differs from ACC2 consensus, and the consensus ACC2 protein is likely functional based on sequence information from multiple tolerant accessions); Green diamond, likely not deleterious (ACC2 variant found in tolerant or high intermediate accession; or ACC1 variant found in single accession not tested for sequence confirmation); Black dot, not deleterious (ACC1 variant found in natural accessions).

Sixth row: Protein domains; 1, Transit peptide; 2, Biotin carboxylase; 3, Biotin carboxyl carrier protein; 4, Central domain; 5, Carboxyltransferase, beta subunit; 6, Carboxyltransferase, alpha subunit; *, Biotin binding site within the BCCP domain.


MEMRALGSSCSTGNGGSAPITLTNI SPWITTVFPSTVKLRSSLRTFKGVS SRVRTFKGVSSTRVLSRTKQQFPLF CFLNPDPISFLDNDVSEAERTVVLP 100
ACCase
ACC2
ACC1
Effect
Domain
DGSVNGAGSVNGYHSDVVPGRN---VAE VNEFCKALGGKRPIHSILVATNGMA AVKFIRSVRTWAYETFGSEKAVKLV AMATPEDMRINAEHIRIADQFVEVP 200
ACCase
ACC2
ACC1
Effect
Domain
GGTNNNNYANVQLIVEMAEVTRVDA VWPGWGHASENPELPDALKEKGIIF LGPPADSMIALGDKIGSSLIAQAAD VPTLPWSGSHVKIPPGRSLVTVPEE 300
ACCase
ACC2
ACC1
Effect
Domain
IYKKACVYTTEEAIASCQVVGYPAM IKASWGGGGKGIRKVHNDDEVRALF KQVQGEVPGSPIFIMKVASQSRHLE VQLLCDQYGNVAALHSRDCSVQRRH 400
ACCase
ACC2
ACC1
Effect
Domain
QKIIEEGPITVAPQETIKKLEQAAR RLAKSVNYVGAATVEYLYSMDTGEY YFLELNPRLQVEHPVTEWIAEVNLP AAQVAVGMGIPLWQIPEIRRFYGME 500
ACCase
ACC2
ACC1
Effect
Domain


HGGGYDSWRKTSVVASPFDFDEAES LRPKGHCVAVRVTSEDPDDGFKPTS GEIQELSFKSKPNMWSYFSVKSGGG IHEFSDSQFGHVFAFGESRSVAIAN 600
d P ag t I ktqr rP vi c I E S tV N R Ss V G g a A i y N ea Rk ACCase
L LN A H Q S ACC2
A fL k q i S rv v a al ACC1
▪ !▪ ■◊! ! ! ▪ ■ !!◊ ◊ ! ! ◊ !! Effect
2222222222222222222222 2222222222222222222222 2222222222222222222222 2222222222222222222222 Domain
MVLALKEIQIRGDIRTNVDYTIDLL HASDYRENKIHTGWLDSRIAMRVRA ERPPWYLSVVGGALYKASTTSSAVV SDYVGYLEKGQIPPKHISLVHSQVS 700
V LS F T E L K ETe Fed t t eL eK t Dtm a C t haa e c e lhs V d L tvfp e ACCase
I W I G Q ACC2
g e d a a M ACC1
! ! ◊ ◊! ◊ ! ! ◊ ◊ ♦ Effect
2222222222222222222222 2222222222222222222----- ---------- ---------- Domain
LNIEGSKYTIDVVRGGSGTYRLRMN NSEVVAEIHTLRDGGLLMQLDGKSH VIYAKEEATGTRLLIDGRTCLLQND HDPSKLMAETPCKLLRYLVSDNSSI 800
IY kft t sSpdS v f G kcEVgVrr s ls tv w Vaa s EkE N Tq RsPS G v E GgHV ACCase
M E S G E I C RS F SV DN V M M ACC2
N N S e a G m IFH n ACC1
◊ ◊♦ ♦ ▪ ◊ ■ ◊◊ ! ! ▪ ♦◊◊ ◊ ◊◊ ◊■ ◊ ◊ ! ▪♦♦ !◊ Effect
---------- ---------- ---------- --333333333333333333333 Domain
DTDTPYAEVEVMKMCMPLISPASGV IHFKLSEGQAMQAGELIAKLDLDDP SAVRKAKPFRGSFPRLGLPTAISGK VHQRCAATLNAARMILAGYDHKVDE 900
AGQ y lAqE vq iKqP AtLe Di i a k kh e t qL e p ivGe p F l iLeN l mk ACCase
R R V P EALH FE R E I L DR ACC2
a M A l S m n eL h K C e ACC1
! ▪◊ ♦ ! ♦ ! ■ ◊ ! ■ !▪◊! ◊◊ ◊ ▪ ◊ ♦◊ !◊◊ Effect
333333333333*333333333333 3333333333333333333344444 4444444444444444444444444 4444444444444444444444444 Domain
VLQDLLNCLDSPELPFLQWQECFAV LATRLPKDLRNMLELKYKEFEIISK TSLTPDFPAKLLKGILEAHLSSCDE KERGSLERLIEPLMSLVKSYEGGRE 1000
tv e mev RD y E imSa hg i qk ekq rklmaryasnit svklce q ak ds atlvr evffmttq vq qr r lr ACCase
H H T L K I F V R K A S F S ACC2
v n I s r s r n M t a N a ACC1
!◊ ◊ ◊ ◊ ! ◊ ▪ !◊ ! ! ◊! ! ♦ ! ◊ ◊ ◊ ! ◊ ▪ ! ◊ ◊ Effect
4444444444444444444444444 4444444444444444444444444 4444444444444444444444444 4444444444444444444444444 Domain


SHARLIVHSLFEEYLSVEELFNDNM LADVIERMRQQYKKDRLKIVDIVLS HQGIIHKNKLVLRLMEQLVYPNPAA YREKLIRFSALNHTNYSQLALKASQ 1100
g ekaV k L q sghy dk lkL e n d ld V sqVsk a lD rp d l tde aeLte esrt KV Re ACCase
TH T Q Y Q N E S P G ACC2
v l l kn I d t e ACC1
◊◊! ◊ ! ! ◊ !! ♦ ◊ !◊ ! ◊ ! ◊ ◊ Effect
444444444444444444444444 444444444444444444444444 444444444444444444444444 444444444444444444444444 Domain
LLEQTKRSELRSNIARSLSELEMFT EAGENMDTPKRKSAISETMENLVSS SLAVEDALVGLFDHSDHTLQRRVVE TYIHRLYQPYVVKESVRMQWHQSGV 1200
v I chLP h QmE svd ys hyet gwd hrepcl nLke d eyt F V pnF y pwV lAAL V VR A rA eL siqy l d pc ACCase
K Q D R I L R K L D M C F IWL I V GF L ACC2
l d n ri d Ca E r d r l ACC1
◊ ! ◊ ◊ ◊ ! ◊ ! !!◊! ◊♦! ◊ ◊ ◊♦ ◊ ◊◊◊ ◊ ! ◊ ◊◊ ! ! ! Effect
444444444444444444444444 444444444444444444444444 444444444444444444444444 444444444444444444444444 Domain
IASWEFLE-HFERKNTGPDDHEISEK GIVAKSSKRKRGTMVIIKSLQFLPS IINASLRETNHSHCEYARAPLSGNM MHIAVVGINNQMSLLQDSGDEDQTQ 1300
v q id lsylv m gssdtp l Rk e v eqrm a af edfee ldea k sp tsdpr pssls v i l a e dd e a ACCase
L G M V II C R K N H R E H I ACC2
I E e m i l NPdt l G H w a T a AM ndS t g F i a ACC1
▪ ▪■ ! ! ! ! ▪▪!! ! ♦ ▪ ! !◊ ◊ ◊ ▪ ! ◊♦♦ !!▪ ! ! ▪ ! ◊ ◊ ◊ ◊ ! Effect
444444444444444444444444 444444444444444444444444 444444444444444444444444 444444444444444444444444 Domain
ERVNKLAKILKEEEVSLTLCSAGVG VISCIIQRDEGRTPMRHSFHWLMEK QYYVEEPLLRHVEPPLSVYLELDKL KGYSNIQYTPSRDRQWHMYSVTDRP 1400
ide rvq nkse lar R R TF g kd ey kyfT Rgpd l e Drii l A AFQ gR sNFnd kpv teN ni l e aKkg ACCase
D A VD K L G Q NV H S ACC2
H ss C sl I l i l t k ACC1
▪ ◊ ◊ !! ◊◊ ▪ ◊ ◊ !! ♦♦ ◊ ! ! ◊◊ ◊ ◊ ! ! ! Effect
444444444444444444444444 444444444444444444444444 444444444444444444444444 444444444444444444444444 Domain
VPIKRMFLRSLVRQTTMNDGFLLQQ GQDYQLSQTVLSMAFTSKCILRSLM NAMEELELNAHNAAMKPDHAHMFLC ILREQQIDDLVPYPRRFEVNAEDEE 1500
tevy F t Aii pp n ts g Lrtki asaEYLi Ea e L l d lD vafn tnvrs cN I NF vfktvimd gq ACCase
N Q KS A IV PQY H N H K ACC2
A a M Yk L li I M v d L A Lf v G ACC1
♦ ◊ ■ ◊! ◊ ▪ ▪! ♦ !! ♦◊ ♦◊◊◊! ! ◊ ◊ ◊ ♦ ♦ ▪! ! ♦ ◊ Effect
444444444444444444444444 444444444444444444444444 444444444444444444444444 444444444444444444444444 Domain


1600 ACCase ACC2 ACC1 Effect Domain
1700 ACCase ACC2 ACC1 Effect Domain
1800 ACCase ACC2 ACC1 Effect Domain
1900 ACCase ACC2 ACC1 Effect Domain
2000 ACCase ACC2 ACC1 Effect Domain


2100 ACCase ACC2 ACC1 Effect Domain
2200 ACCase ACC2 ACC1 Effect Domain
2300 ACCase ACC2 ACC1 Effect Domain
2355 ACCase ACC2 ACC1 Effect Domain

APPENDIX K: Informative Variants that Alter Conserved Residues in Eukaryotic, Homomeric Acetyl-CoA Carboxylases

This appendix lists the details of informative variants and residues in the ACC2 protein sequence. Included data are the position in the ACC2 protein sequence; the variant at that position; the variant type; the organism the variant is found in; the current classification of the variant based on all information known; the allele strength; the locus and domain where the residue is found; the percent conservation of the consensus amino acid at each residue for three alignments: (1) the original multi-kingdom alignment of 20 eukaryotic sequences, (2) the alignment of 139 plant sequences, and (3) the multi-kingdom alignment of 667 eukaryotic sequences; the genotype each variant is found in; the residue location in the source organism; relevant references, and additional notes. Adapted from Parker et al. (2016).

Footnotes for the title row of the following table are described below:

[a] The first residue (e.g. "G" in G135E) is found in the consensus sequence among the 857 accessions; the second is the variant.

[b] D, deleterious to protein function; LD, likely deleterious; PD, possibly deleterious; VUS, variant of unknown significance; LND, likely not deleterious; ND, not deleterious.

[c] BC, biotin carboxylase; BCCP, biotin carboxyl carrier protein; CT-Beta, carboxyltransferase-beta subunit; CT-Alpha, carboxyltransferase-alpha subunit.

Residue (ACC2)	Variant (ACC2) [a]	Variant Type	Organism	Variant Impact [b]	Allele Strength	Locus	Domain [c]	Conservation (%)			Genotype	Source AA Residue	Reference	Additional Notes
								Original (20)	Plant (139)	MUSCLE (667)				
135	G135E	Missense	Arabidopsis	LD	Strong	*ACC2*	(BC)	95	100.0	95.7	Sav-0	Same	Parker et al. (2016)	Unique to Sav-0 Accession
153	K153E	Site Directed	Yeast	D	Strong	*ACC1*	Dimer Interface	100	100.0	96.9	K73E	73	Wei and Tong (2015)	Loss of Enzyme Activity *in vitro*
156	R156E	Site Directed	Yeast	D	Strong	*ACC1*	Dimer Interface	100	99.3	93.6	R76E	76	Wei and Tong (2015)	Loss of Enzyme Activity *in vitro;* Soraphen A Interaction Site
188	E188K	Missense	Arabidopsis	D	Weak	*ACC1*	BC	90	99.3	78.0	*gsd1*	86	Lü et al. (2011)	Vegetative Phenotype
193	A193V	Missense	Arabidopsis	ND	Normal	*ACC1*	BC	100	99.3	98.4	Melni-2	91	Parker et al. (2016)	Maintained in Natural Populations
219	219	Splicing	Arabidopsis	D	Strong	*ACC1*	BC	NA	NA	NA	*gk-U413; gk-sc*	114	Kajiwara et al. (2004)	Embryo Defective; Updated Location
333	G333D	Missense	Arabidopsis	D	Strong	*ACC1*	BC	100	97.1	99.4	*acc1-3*	231	Kajiwara et al. (2004)	Embryo Defective; Seeds Unavailable
363	F363L	Missense	Arabidopsis	PD	Some Function	*ACC2*	BC	100	97.1	99.3	Sei-0	Same	Parker et al. (2016)	
376	V376A	Missense	Arabidopsis	PD	Some Function	*ACC2*	BC	100	100.0	100.0	Col-0	Same	Parker et al. (2016)	
383	Y383H	Missense	Arabidopsis	ND	Normal	*ACC1*	BC	80	71.9	90.6	Consensus		Parker et al. (2016)	ACC1 Consensus Differs from ACC2
397	Q397X	Nonsense	Drosophila	D	Strong	*ACC*	BC	NA	NA	NA	*Acc*^{1}	359	Sasamura et al. (2013)	Lethal
402.3	x	x	Arabidopsis	x	x	*x*	Large Intron	x	x	x	x	x	x	Large Intron
404	I404K	Missense	Arabidopsis	LD	Strong	*ACC2*	BC	95	100.0	94.8	Knox-18 Group	Same	Parker et al. (2016)	See Others in Group
406	E406K	Missense	Arabidopsis	D	Weak	*ACC1*	BC	100	100.0	100.0	*sfr1*	304	Amid et al. (2012)	Vegetative Phenotype
443	Y443C	Missense	Arabidopsis	VUS	Uncertain	*ACC2*	BC	95	92.8	94.0	Etna-2	Same	This dissertation	
456	N456I	Missense	Drosophila	D	Strong	*ACC*	BC	100	100.0	100.0	*Acc*^{2}	417	Sasamura et al. (2013)	Lethal
474	L474F	Missense	Arabidopsis	PD	Some Function	*ACC2*	BC	100	97.8	94.5	Chi-0	Same	Parker et al. (2016)	

475	P475L	Missense	Arabidopsis	LND	Normal	*ACC2*	BC	100	99.3	99.7	Lm-2	Same	Parker et al. (2016)	
478	Q478K	Missense	Arabidopsis	LND	Normal	*ACC2*	BC	100	99.3	97.6	Uod-1	Same	Parker et al. (2016)	
493	I493L	Missense	Arabidopsis	ND	Normal	*ACC1*	BC	100	99.3	94.6	Multiple	391	Parker et al. (2016)	Maintained in Natural Populations
494	R494G	Missense	Arabidopsis	PD	Some Function	*ACC2*	BC	100	99.3	99.9	Ip-Pal-0	Same	Parker et al. (2016)	
520	F520L	Missense	Arabidopsis	PD	Some Function	*ACC2*	BC	40	82.7	94.0	In-0	Same	Parker et al. (2016)	
528	P528S	Missense	Arabidopsis	ND	Normal	*ACC1*	BC	90	99.3	90.1	Multiple	426	Parker et al. (2016)	Maintained in Natural Populations
538	T538A	Missense	Arabidopsis	PD	Some Function	*ACC2*	BC	100	99.3	99.9	Ip-Tor-1	Same	Parker et al. (2016)	
564	M564V	Missense	*C. elegans*	(D)	Weak		BC	95 (V)	93.5 (V)	93.3 (V)	*ye60* (A471V)	471	Rapppleye et al. (2003)	Temperature Sensitive
565	W565A	Site Directed	Yeast	D	Strong	*ACC1*	Dimer Interface	100	97.1	99.3	W487A	487	Wei and Tong (2015)	Loss of Enzyme Activity *in vitro*
565	W565X	Nonsense	Arabidopsis	D	Strong	*ACC1*	BC	NA	NA	NA	*emb22*	463	Kajiwara et al. (2004)	Embryo Defective
572	See Text	Splicing	Arabidopsis	D	Strong	*ACC2*	Intron 10	NA	NA	NA	Gn-1; "Gn2-3"	Same	Parker et al. (2016)	Results in Frameshift
668	S668S	Missense	Yeast	(D)	Weak	*ACC1*		35 (S)	68.4 (S)	33.0 (H)	*acc1*^{ts} (F>S)		Schneiter et al. (2000)	
686	Q686R	Site Directed	Yeast	LND	Normal	*ACC1*		95	99.3	96.7	Q608R	608	Wei and Tong (2015)	Functional Enzyme *in vitro*
725	N725S	Missense	Arabidopsis	LND	Normal	*ACC2*		100	95.7	96.9	Pog-0	Same	Parker et al. (2016)	
734	H734E	Site Directed	Yeast	LND	Normal	*ACC1*		70 (H)	69.8 (H)	44.2 (R)	R656E	656	Wei and Tong (2015)	Functional Enzyme *in vitro*
739	G739E	Missense	Arabidopsis	PD	Some Function	*ACC2*		95	97.8	95.2	Wa-1	Same	Parker et al. (2016)	
753	Y753X	Nonsense	Arabidopsis	D	Strong	*ACC2*		NA	NA	NA	Kb-0; Kl-5	Same	Parker et al. (2016)	
762	R762C	Missense	Arabidopsis	ND	Normal	*ACC2*		100	97.1	96.6	Tsu-0; Tu-0	Same	Parker et al. (2016)	
774	See Fig. S2	Deletion	Arabidopsis	D	Strong	*ACC2*	Intron 17; Exon 18	NA	NA	NA	Ip-Ber-0	Same	Parker et al. (2016)	23 bp Deletion; Defective Transcripts

777	D777N	Missense	Arabidopsis	PD	(Some Function)	*ACC2*	(BCCP)	95	100.0	97.2	Leska-1-44; Sei-0	Same	Parker et al. (2016)	
794	V794I	Missense	Arabidopsis	ND	Normal	*ACC1*	BCCP	95	84.9	85.6	Multiple	692	Parker et al. (2016)	Maintained in Natural Populations
813	K813R	Site Directed	Yeast	D	Strong	*ACC1*	Biotin Binding	100	100.0	100.0	Biotin Binding	735	Schneiter et al. (1996)	Site-Directed Mutagenesis
833	G833R	Missense	Arabidopsis	PD	Some Function	*ACC2*	BCCP	100	98.6	99.3	Dja-1	Same	Parker et al. (2016)	
847	L847P	Missense	Arabidopsis	PD	Some Function	*ACC2*	Central	100	100.0	96.0	WAR	Same	Parker et al. (2016)	
865	R865X	Nonsense	Arabidopsis	D	Strong	*ACC2*	Central	NA	NA	NA	"Nossen"	Same	Parker et al. (2014)	
901	See Text	Splicing	Arabidopsis	D	Strong	*ACC2*	Intron 19	NA	NA	NA	Wl-0	Same	Parker et al. (2016)	Results in Frameshift
955	[955]	Insertion	Arabidopsis	D	Strong	*ACC2*	Exon 21	NA	NA	NA	*acc2-2*	Same	Salk Insertion	T-DNA Insertion Mutant
1171	1171fs	Frameshift	Arabidopsis	D	Strong	*ACC2*	Central	NA	NA	NA	Ip-Alo-0; Ip-Vin-0	Same	Parker et al. (2016)	
1206	F1206L	Missense	Arabidopsis	LD	Moderate	*ACC2*	Central	85	100	96.3	Aitba-1	Same	Parker et al. (2016)	
1225	K1225X	Nonsense	Arabidopsis	D	Strong	*ACC2*	Central	NA	NA	NA	Blh1-1	Same	Parker et al. (2016)	
1229	[1229]	Insertion	Arabidopsis	D	Strong	*ACC2*	Exon 27	NA	NA	NA	*acc2-1*	Same	Salk Insertion	T-DNA Insertion Mutant
1355	E1355G	Missense	Arabidopsis	VUS; LND	Uncertain; (Normal)	*ACC2*	Central	100	100.0	98.7	Knox-18 Group; (Si-0; Ema-1)	Same	Parker et al. (2016)	
1376; 1377	K1376R; Δ1377	Deletion	Arabidopsis	VUS	Uncertain	*ACC2*	Central	35 20	80.6 43.9	Low Low	Qar-8a	Same	This dissertation	
1405	R1405Q	Missense	Arabidopsis	PD	Some Function	*ACC2*	Central	100	100.0	96.1	Db-1	Same	Parker et al. (2016)	
1479	Δ1479	Deletion	Arabidopsis	PD	Some Function	*ACC2*	Central	45 (E)	92.1 (E)	19.3 (E)	Ip-Voz-0	Same	Parker et al. (2016)	Arabidopsis ACC2: Glu
1562	See Fig. S1	Splicing	Arabidopsis	D	Strong	*ACC2*	Intron 29	NA	NA	NA	Spro-2; Ste-2; Ste-3; Vimmerby	Same	Parker et al. (2016)	Variety of Defective Transcripts
1603	K1603Q	Missense	Arabidopsis	ND	Normal	*ACC1*		80	88.5	91.0	Consensus		Parker et al. (2016)	ACC1 Consensus Differs from ACC2

1621	x	x	Arabidopsis	x	x	*x*	Start of Large Exon	x	x	x	x	x	TAIR	Start of Large Exon (# 31 of 32)
1623	E1623E	Missense	*C. elegans*		Weak		CT-Beta	35 (E)	87.1 (E)	61.3 (R)	*ye162* (G1351E)	1351	Rapppleye et al. (2003)	
1689	E1689K;G	Missense	Arabidopsis	D	Strong	*ACC1; ACC2*	CT-Beta	100	99.3	97.0	*pas3-1* (E>K); Ts-1 (E>G)	1588; Same	Baud et al. (2004); Parker et al. (2016)	Embryo Defective (*acc1*); Spectinomycin Sensitive (*acc2*)
1739	S1739C	Missense	Arabidopsis	PD	Some Function	*ACC2*	CT-Beta	90	86.3	94.3	CYR	Same	Parker et al. (2016)	Conserved Residue in CoA Binding Pocket
1766	G1766D	Missense	Arabidopsis	LND	Normal	*ACC2*	CT-Beta	100	99.3	100	Pog-0	Same	Parker et al. (2016)	
1794	G1794A	Missense	Arabidopsis	ND	Normal	*ACC1*	CT-Beta	95	100.0	94.8	Multiple	1693	Parker et al. (2016)	Maintained in Natural Populations
1815	I > L,V,A,T	Missense	Resistant Grasses	ND	Normal	*ACC2*	CT-Beta	90 (L)	74.8 (L)	94.3 (L)	Plastid ACCase	1781	Kaundun (2014) GenBank: AJ310767	Herbicide Resistant Grasses; Arabidopsis ACC2: Leu
1821	I1821V	Missense	Arabidopsis	PD	Some Function	*ACC2*	CT-Beta	100	100.0	98.2	MNF-Che-2	Same	Parker et al. (2016)	
1834	T1834S	Missense	Arabidopsis	PD	Some Function	*ACC2*	CT-Beta	100	100.0	99.4	Nemrut-1	Same	Parker et al. (2016)	
1854	[1854]	Insertion	Arabidopsis	D	Strong	*ACC1*	CT-Beta	NA	NA	NA	*acc1-1*	1753	Baud et al. (2004)	
1878	R1878X	Nonsense	Yeast	D		*ACC1*	CT-Beta	NA	NA	NA	$Acc1^{C\text{-}term}$		Schneiter et al. (2000)	
1883	S1883T	Missense	Arabidopsis	PD	Some Function	*ACC2*	CT-Beta	100	100.0	97.0	Several	Same	Parker et al. (2016)	
1888	G1888S	Missense	Arabidopsis	D	Strong	*ACC1*	CT-Beta	100	100	99.4	*pas3-2*	1787	Baud et al. (2004)	Embryo Defective
1889	G1889A	Missense	Yeast	(D)	Weak	*ACC1*	CT-Beta	100	100	99.4	$Acc1^{cs}$		Schneiter et al. (2000)	
1890	P1890C	Missense	*C. elegans*	(D)	Weak		CT-Beta	45 (P)	97.1 (P)	42.1 (T)	*ye180*		Rappleye et al. (2003)	
1897	G1897S	Missense	Arabidopsis	PD	Some Function	*ACC2*	CT-Alpha	100	100.0	99.4	Sch1-7; WalHaesB4	Same	Parker et al. (2016)	

1902	T1902K	Missense	Arabidopsis	LD	Strong	*ACC2*	CT-Alpha	95	100.0	87.6	Knox-18 Group	Same	Parker et al. (2016)	See Others in Group
1968	G1968E	Missense	Arabidopsis	D	Moderate	*ACC1*	CT-Alpha	100	99.3	100	*gk-101*	1867	Kajiwara et al. (2004)	Embryo Defective; Seeds Unavailable
2013	P2013L	Missense	Arabidopsis	PD	Some Function	*ACC2*	CT-Alpha	95	97.1	98.5	Balan-1	Same	Parker et al. (2016)	
2014	A2014E	Missense	Arabidopsis	PD	Some Function	*ACC2*	CT-Alpha	100	97.8	99.0	App1-16	Same	Parker et al. (2016)	
2020	2020fs	Frameshift	Arabidopsis	D	Strong	*ACC2*	CT-Alpha	NA	NA	NA	Lu4-2; Lu3-30	Same	Parker et al. (2016)	
2033	W > C,L,S	Missense	Resistant Grasses	ND	Normal	*ACC2*	CT-Alpha	100	98.6	98.5	Plastid ACCase	1999	Kaundun (2014) GenBank: AJ310767	Herbicide Resistant Grasses; Arabidopsis ACC2: Trp
2059	A2059V	Missense	Arabidopsis	VUS	Uncertain	*ACC2*	CT-Alpha	100	98.6	98.2	Grivo-1	Same	This dissertation	
2061	W > C	Missense	Resistant Grasses	ND	Normal	*ACC2*	CT-Alpha	100	98.6	97.3	Plastid ACCase	2027	Kaundun (2014) GenBank: AJ310767	Herbicide Resistant Grasses; Arabidopsis ACC2: Trp
2098	P2098S	Missense	Arabidopsis	VUS	Some Function	*ACC2*	CT-Alpha	100	97.8	93.0	Hod	Same	Parker et al. (2016)	Nonsense Mutation Also Present
2112	D > G	Missense	Resistant Grasses	ND	Normal	*ACC2*	CT-Alpha	100	97.8	98.5	Plastid ACCase	2078	Kaundun (2014) GenBank: AJ310767	Herbicide Resistant Grasses; Arabidopsis ACC2: Asp
2115	I2115R	Missense	Arabidopsis	PD	Some Function	*ACC2*	CT-Alpha	100	97.8	98.2	Iasi-1	Same	Parker et al. (2016)	
2122	C > R	Missense	Resistant Grasses	ND	Normal	*ACC2*	CT-Alpha	80 (M)	70.5 (M)	81.9 (M)	Plastid ACCase	2088	Kaundun (2014) GenBank: AJ310767	Herbicide Resistant Grasses; Arabidopsis ACC2: Met
2130	G > A,S	Missense	Resistant Grasses	ND	Normal	*ACC2*	CT-Alpha	90	96.4	80.1	Plastid ACCase	2096	Kaundun (2014) GenBank: AJ310767	Herbicide Resistant Grasses; Arabidopsis ACC2: Gly
2207	H2207Q	Missense	Arabidopsis	PD	Some Function	*ACC2*	CT-Alpha	100	98.6	98.1	Ip-Lso-0	Same	Parker et al. (2016)	

2208	[2208]	Insertion	Arabidopsis	D	Strong	*ACC1*	CT-Alpha	NA	NA	NA	*acc1-2*	2107	Baud et al. (2004)	
2325	Q2325X	Nonsense	Arabidopsis	LD	Some Function	*ACC2*		NA	NA	NA	Hod	Same	Parker et al. (2016)	
2337	x	x	Arabidopsis	x	x	*x*	End of Large Exon	x	x	x	x	x	TAIR	End of Large Exon (# 31 of 32)
2355	x	x	Arabidopsis	x	x	*x*	End of Protein	x	x	x	x	x	TAIR	End of ACC2 Protein

APPENDIX L: ACC2 Variants that Alter Conserved Residues in Sequenced Arabidopsis Accessions

This appendix lists the details of all variants among Arabidopsis accessions that differ than the ACC2 consensus protein sequence. Included data are the position number and amino acid substitution for each variant, the percent conservation of the consensus amino acid based on our multi-kingdom alignment of 667 eukaryotic ACCase sequences, the number of accessions with each variant, and data on the accessions containing the variant that have been analyzed on spectinomycin: (1) name of the accessions; (2) whether the variant was confirmed in the accession or not; (3) the number of seedlings screened on spectinomycin; and (3) the category and score of the spectinomycin response. Adapted from Parker et al. (2016).

Footnotes for the following table are described below:

[a] The first residue (e.g. "G" in G135E) is found in the consensus sequence among the accessions; the second is the variant.

[b] Percentage of 667 aligned homomeric ACCase sequences with the accession consensus residue. Red, ≥ 99% conserved.

[c] BC, Biotin carboxylase; BCCP, Biotin carboxyl carrier protein; CT, carboxyltransferase.

[d] Accessions with the same variant but a more sensitive or problematic seedling response are excluded to highlight the most tolerant responses observed with the variant present.

[e] The Sav-0 variant was uncovered by sequencing the *ACC2* cDNA.

[f] Intermediate responses: Aa-0; Hsm; Kyoto; Rag1-1; Ws-2.

[g] Intermediate responses: Boot-1; Col-0; Ga-0; Hi-0; Kn-0; Ler-1; NFA-8; Pi-0; Tscha-1; Van-0.

[h] Intermediate responses: Ber; CON-7; Dja-1; Est; Gy-0; Nie1-2; Pla-0; Sch1-7; Wa-1; WalHaesB4.

[i] Intermediate responses: Fei-0; Kin-0; Seattle-0; Sq-8.

[j] Intermediate responses: Boot-1; Col-0; Kn-0; Pi-0; Van-0.

[k] Intermediate responses: Dra3-1; Kni-1; Pna-17; Spr1-2.

Variant [a]	Conservation (%) [b]	Protein Domain [c]	1001 Genomes Accessions with Predicted Variant	Accessions Evaluated on Spectinomycin [d]	Variant Confirmed	Seedlings Classified	Spectinomycin Response	
							Category	Score
G135E	95.7	(BC)	0	Sav-0 [e]	Yes	275	Hypersensitive	1.2
F363L	99.3	BC	5	Sei-0	Yes	56	Intermediate	6.9
V376A	100.0	BC	12	Col-0	Yes	287	Intermediate	5.6
I404K	94.8	BC	20	Knox-18 Group	Yes	-	Hypersensitive; Sensitive	-
Y443C	94.0	BC	1	Etna-2	Yes	111	Sensitive	1.9
L474F	94.5	BC	1	Chi-0	Yes	75	Intermediate	5.1
P475L	99.7	BC	1	Lm-2	Yes	70	Tolerant	8.3
Q478K	97.6	BC	28	Multiple [d,f]	Assumed	126	Intermediate	6.0
				Uod-1	Yes	78	Tolerant	8.5
R494G	99.9	BC	1	Ip-Pal-0	Yes	51	Intermediate	6.1
F520L	94.0	BC	2	In-0	Yes	84	Intermediate	4.8
T538A	99.9	BC	1	Ip-Tor-1	Yes	46	Low Intermediate	4.1
V618I	90.9	BC	9	Ip-Cum-1 Ped-0	Yes	129 56	Sensitive	2.3 3.1
				Ip-Gua-1 Ip-Hom-4	Yes	81 94	Intermediate	6.3 4.5
N725S	96.9		44	Multiple [d,g]	Assumed	496	Intermediate	5.8
				Pog-0	Yes	60	Tolerant	8.6
G739E	95.2		1	Wa-1	Yes	40	Intermediate	5.3
R762C	96.6		6	Mh-0	Not Tested	28	Intermediate	6.2
				Tsu-0 Tu-0	Yes	490 84	Tolerant	8.8 9.4
D777N	97.2		4	Can-0	NO	96	Intermediate	4.7
G833R	99.3	BCCP	3	Dja-1	Yes	52	Intermediate	4.6
L847P	96.0		1	WAR	Yes	39	Low Intermediate	3.7
F1206L	96.3		1	Aitba-1	Yes	53	Sensitive	2.8
E1355G	98.7		116	Multiple [d,h]	Assumed	625	Intermediate	5.1
				Ema-1; Si-0	Yes	124	High Intermediate	8.2
R1405Q	96.1		1	Db-1	Yes	75	Intermediate	7.2
Y1594H	97.9		6	None	Not Tested	None	Not Tested	
E1689G	97.0	CT-β	1	Ts-1	Yes	70	Sensitive	2.5

S1739C	94.3	CT-β	12	Multiple [d,i]	Assumed	97	Intermediate	5.6
				CYR	Yes	76	High Intermediate	7.8
G1766D	97.6	CT-β	39	Multiple [d,j]	Assumed	380	Intermediate	5.7
				Pog-0	Yes	60	Tolerant	8.6
I1821V	98.2	CT-β	1	MNF-Che-2	Yes	53	Intermediate	4.3
T1834S	99.4	CT-β	2	Nemrut-1	Yes	75	Intermediate	4.2
S1883T	97.0	CT-β	8	Multiple [d,k]	Assumed	211	Intermediate	6.2
G1897S	99.4	CT-α	2	Sch1-7 WalhaesB4	Yes	70 39	Intermediate	4.8 5.6
P2013L	98.5	CT-α	3	Balan-1	Yes	52	Intermediate	6.2
A2014E	99.0	CT-α	1	Appl-16	Yes	54	Intermediate	5.6
A2059V	98.2	CT-α	1	Grivo-1	Yes	73	Sensitive	2.0
P2098S	93.0	CT-α	2	Hod	Yes	72	Intermediate	5.3
I2115R	98.2	CT-α	1	Iasi-1	Yes	44	Intermediate	5.0
H2207Q	98.1	CT-α	1	Ip-Lso-0	Yes	55	Intermediate	7.5

VITA

Type Nicole Bryant Parker

Candidate for the Degree of

Doctor of Philosophy

Thesis: NATURAL VARIATION IN SENSITIVITY TO A LOSS OF CHLOROPLAST TRANSLATION IN ARABIDOPSIS

Major Field: Plant Science

Biographical:

Education:

- Completed the requirements for the Doctor of Philosophy in Plant Science at Oklahoma State University, Stillwater, Oklahoma in December 2017
- Completed the requirements for the Bachelor of Science in Botany at Oklahoma State University, Stillwater, Oklahoma in May 2011.
- Completed the requirements for the Bachelor of Science in Microbiology at Oklahoma State University, Stillwater, Oklahoma in May 2011.

Experience:

- Various Teaching Assistantships from August 2011 to May 2017 including Plant Anatomy, Plant Biology, Introductory Biology, and Genetics.
- Three publications as lead author in Plant Physiology. Six poster presentations at local and international conferences. Five oral presentations at local seminars and international conferences.
- Doctoral Research, Department of Botany, Oklahoma State University, Stillwater, Oklahoma from August 2011 to May 2017
- Undergraduate Honors Research, Department of Botany, Oklahoma State University, Stillwater, Oklahoma from August 2009 to July 2011.
- Student Lab Technician, Department of Botany, Oklahoma State University, Stillwater, Oklahoma from January 2009 to August 2010.

Professional Memberships:

- Oklahoma State University Botanical Society
- Botany Graduate Student Association
- Graduate and Professional Student Government Association
- Oklahoma Academy of Science
- Golden Key International Honour's Society
- Phi Kappa Phi Honors Society